

A CRITICAL STUDY OF THE COHERENCE OF CRITERIAL REASONING.

PAUL TOMASSI



PH.D
UNIVERSITY OF EDINBURGH
1989.

ABSTRACT

The thesis focusses on the question of the coherence of the notion of a "criterion" which has enjoyed considerable currency in the contemporary Philosophy of Language and Mind. The text is divided into three parts, each of which deals with a particular aspect of the notion of a criterion, or of theories of criteria. Part One consists of Chapters 1, 2 and 3 and explores the question of the origins of the notion and of theories concerning it. Wittgenstein's later work is widely recognized as the notion's point of origin and for that reason I examine Wittgenstein's use of the term "criterion" and the philosophy of language which is the proper context of the notion as employed by that author. I then present my own account of the significance of the notion to Wittgenstein, and the role of the notion in his later thought. Part Two consists of Chapters 4 and 5 which examine the place of the notion in the contemporary Philosophy of Language and offer a critical account of much recent work in that field. Finally, in Part Three, I concentrate on the question of the coherence of the notion. The sixth Chapter deals mainly with epistemological issues of relevance but also involves some formal work together with some discussion of the ramifications of the notion, correctly understood, in the Philosophy of Mind. Chapter Seven suggests a formal, logical framework in which criterial reasoning patterns can be usefully mapped and also attempts to demonstrate an inherent flexibility of format which might allow for contrastive comparisons with other modes of reasoning.

The following thesis consists wholly of my own work and thoughts on the subjects discussed therein.

Paul Tomassi

CONTENTS

PART ONE: THE ORIGINS OF THE THEORY OF CRITERIA.

CHAPTER ONE..... 1.

- INTRODUCTION
- I COMMENTARY
- II EXEGESIS
- III SOCIOLOGY
- IV USE
- V CONSISTENCY

CHAPTER TWO..... 31.

- INTRODUCTION: A PARADOX AT THE HEART OF WITTGENSTEIN'S
LATER PHILOSOPHY OF MIND
- I THE LANGUAGE OF SENSATION
- II EXPERIENCE AND MEANING
- II BEHAVIOUR AND SENSATION
- IV CRITERIA AND EMOTION

CHAPTER THREE..... 60.

- INTRODUCTION
- I THEORY OF MEANING
- II OTHER MINDS
- III ON CERTAINTY

PART TWO: THE ROLE AND PLACE OF THE THEORY OF CRITERIA IN THE CONTEMPORARY PHILOSOPHY OF LANGUAGE AND MIND

CHAPTER FOUR..... 87.

- INTRODUCTION
- I THE PHILOSOPHY OF LANGUAGE
- II THE SCIENCE OF LINGUISTICS: A HYPOTHETICO-DEDUCTIVE MODEL
- III THE MACHINE IN THE GHOST
- IV WHAT IS A THEORY OF MEANING ?
- V A CRITICAL APPRAISAL
- VI AN ALTERNATIVE SOLUTION
- VII A COMPARISON: LANGUAGE AND STRUCTURE IN THE MODERN CONTINENTAL
TRADITION

CHAPTER FIVE..... 136.

- INTRODUCTION
- I SCEPTICISM, RULES AND LANGUAGE
- II CRITERIA AND INTERNAL RELATIONS
- III FIRST THOUGHTS

PART THREE: ON THE COHERENCE OF CRITERIAL REASONING

CHAPTER SIX.....	156.
------------------	------

	INTRODUCTION
I	DEFEASIBILITY AND KNOWLEDGE
II	JUSTIFICATION AND BELIEF
III	DEFEASIBILITY AND THE MEANING FEATURE
IV	CONVENTION AND CONTINGENCY

CHAPTER SEVEN.....	205.
--------------------	------

	INTRODUCTION
I	THE DIALOGUE SYSTEM
II	RULES AND ACTS
III	DOUBT AND GROUNDS: TWO CASES
IV	GROUNDING DOUBT

BIBLIOGRAPHY.

CHAPTER ONE

INTRODUCTION

The concept of a criterion makes its full philosophical debut in the works of Wittgenstein in the early pages (24-25) of the Blue Book [1933-34]. Nowhere else is the concept given such overt, sustained and self-conscious consideration though it appears in a wide variety of later contexts. In fact the term first occurs in a number of passages (III 23, XVII 181, 186, XX 206, 215) in the Philosophical Remarks [1929-30], which would place its exact origin some four years earlier. As the primary focus of contemporary debate, however, the Blue Book passage remains very much its beginning. For that reason I sacrifice exact textual chronology in order to attain an ubersicht of commentary.

I COMMENTARY

Baker and Hacker, who for many years championed a criteriological interpretation of the later Wittgenstein, originally shared a particular opinion of these passages. Baker [1974, p.159] referred to "the ambiguity of 'criteria' in the Blue Book" which resulted from "the notorious passage which introduces criteria" and concluded that "the term must be judged ambiguous in the Blue Book". Similarly, Hacker [1972, p.285] spoke of "inconsistencies in the Blue Book's use of the term and maintained that "this notorious passage (Bl. B p.24-25) is misleading". At that time both authors agreed that with the exception of the offending introduction there did exist a clear and univocal later use with which the introduction

could not sit comfortably. If this account is impartial exegesis and therefore valid commentary, then, as Nietzsche [1973, p. 131] has asserted: "the disinterested act is an interesting and interested act".

In the most recent of Hacker's commentaries, the revised edition of "Insight and Illusion" [1985], that attitude is significantly moderated. In this case a textual 'ambiguity' is noted, though the point is relegated to a footnote [p.316]. The ambiguity in question concerns whether or not descriptions of the satisfaction of P's criterial circumstances actually entail that P is true. The problem arises because, to Hacker, the two examples of criteria which Wittgenstein presents in the Blue Book p.24-25 each seem to suggest a different answer to that question. This is clearly the cause of the original claim. That is, at least one of the examples in the introductory passages was felt to portray descriptions of the satisfaction of P's criterial circumstances as entailing P while in many later uses descriptions of the satisfaction of criteria for P appeared to fall short of entailing that P.

While I have no hermeneutic axe to grind as regards defeasibility, it does seem to me that if it is not immediately obvious that the relation was specifically intended not to be entailing it is equally not obvious that it was intended purely as one of entailment. In fact I shall argue both that Wittgenstein did not distinguish between criteria, or between criteria and symptoms, in terms of the notion of entailment and that these passages present a mish-mash of ideas not all of which were later taken up by Wittgenstein. What is patently obvious is that explicit concern and discussion of defeasibility and criterial relations does not feature in any of the passages in

question and that is consistently the case as regards later uses. I return to the issue of consistency below in section V.

In sharp commentorial contrast to this type of interpretative line John Cook [1969, p.134-5] has proposed an alternative exegesis of these passages whereby Wittgenstein is considered to have exposed the concept of a criterion which he had just introduced as being "a piece of confusion" and ultimately "a bogus notion". For Cook the whole exercise forms perhaps the earliest example of what was later to be termed Wittgenstein's attack on essentialism. Apparently then, Wittgenstein introduced the term only to explode it as a myth, a conceptual confusion. Cook [1969] presupposes that reasoning criterially is on a philosophical par with Socratic Dialectic. According to Cook [1969], in identifying the criterion for an expression we not only discover the real meaning of that expression but identify the single form, the "common element", which every proper case of application manifests. Certainly, such classically rationalist epistemology and its metaphysically essentialist underpinnings are anathema to the spirit of Wittgenstein's later work. No doubt he would have wholeheartedly opposed it, that is, if he had characterized the concept of a criterion as Cook has alleged. During his account of the matter Cook [op. cit] does acknowledge that: "the argument is not made explicit in the Blue Book although the conclusion is stated on page 25". This is explained by the fact that the Blue Book was originally a supplement to a series of lectures and therefore was intended only for those attending the lectures which accompanied it. Against this interpretation I argue that while Baker and Hacker were visibly too extreme in attributing

a full-blown criterial theory to Wittgenstein, Cook goes too far in the opposite direction if, as appears to be the case, he is allowing no positive use of the term at all.

II EXEGESIS

Over pages 24-25 of the Blue Book Wittgenstein deploys the terms "criteria"/"criterion" nine times in total. It may prove valuable to examine each of these occurrences and their contexts so as to reconstruct the basic framework of Wittgenstein's argument here.

Firstly, a criterion for an expression is identified as an aspect of that expression's grammar, sense or meaning and to explain one's criterion for an assertion containing that expression is considered tantamount to giving an explanation of the meaning of that expression:

"..to explain my criterion for another person's having toothache is to give a grammatical explanation about the word 'toothache' and in this sense, an explanation concerning the meaning of the word 'toothache'."
(B1B.p.24)

Secondly, criteria are alleged to function as means of identification which co-ordinate experience and utterance. Criteria are construed as phenomena which are themselves coincident with other phenomena and are described as being in some sense conventional, I discuss below just in what sense that is so.

"Suppose that by observation I found that in certain cases whenever these first criteria told me a person had toothache, a red patch appeared on the person's cheek. Supposing I now said to someone "I see A has toothache, he's got a red patch on his cheek." He may ask me "How do you know A has toothache when you see a red patch?" I should then point out that certain phenomena had always coincided with the appearance of the red patch. Now one may go on and ask: "How do you know he has got toothache when he holds his cheek?" The answer to this might be, " I say HE has toothache when he holds his cheek because I hold my cheek when I have toothache." But what if we went on asking..

You will be at a loss to answer ..and find that here we strike rock bottom, that is we have come down to conventions."
(B1B.p.24)

The third use introduces "criteria" as a term to be contrasted with the term "symptom":

"Let us introduce two antithetical terms in order to avoid certain elementary confusions: To the question 'How do you know that so-and-so is the case?', we sometimes answer by giving 'criteria' and sometimes by giving 'symptoms'."

The introduction of new terminology is in keeping with the promise on page 23 that not only will we examine particular uses so as to 'clear the matter up' but:

"We shall also try to construct new notations, in order to break the spell of those we are accustomed to."

The fourth and fifth uses occur in a passage where it is made clear that medical science, in the sense of the community of doctors and medical scientists, is that authority which decides on the defining criterion for the term 'angina' and that the criterion contrasts with 'symptoms':

"If medical science calls angina an inflammation caused by a particular bacillus, and we ask in a particular case 'why do you say this man has got angina?' then the answer 'I have found the bacillus so-and-so in his blood' gives us the criterion, or what we may call the defining criterion of angina. If on the other hand the answer was, 'his throat was inflamed', this might give us a symptom of angina."

The sixth use occurs when Wittgenstein explains a 'symptom' as a phenomenon which is coincident with the presence of a criterion, and reiterates and emphasises the fact that the defining criterion will itself be a phenomenon:

"I call a 'symptom' a phenomenon of which experience has taught us that it coincided, in some way or another with the phenomenon which is our defining criterion. Then to say 'A man has angina if this bacillus is found in him' is a tautology or it is a loose way of stating the

definition of 'angina'. But to say, 'A man has angina whenever he has an inflamed throat' is to make an hypothesis."

The seventh and eighth uses point up the fundamental interchangeability of criteria and symptoms and emphasise the fact that although we can decide to give a word a use on the basis of criteria selecting a criterion from a given set of phenomena may be wholly arbitrary or ad hoc.

"In practice if you were asked which phenomenon is the defining criterion and which is a symptom you would in most cases be unable to answer this question except by making an arbitrary decision ad hoc. It may be practical to define a word by taking one phenomenon as the defining criterion but we shall easily be persuaded to define the word by means of what, according to our first use was a symptom."
(BlB.p.25)

It is tempting to describe the process of deciding on a criterion as essentially a matter of convention, particularly in the case of angina. If we did so describe the second example we might then argue that both examples of criteria are identical in presenting criteria as conventional. But there are at least two senses of "convention" at work here. Firstly, Wittgenstein speaks of 'coming down to conventions' and secondly, we could describe doctors as deciding 'by convention' on a criterion for angina. The two uses, and therefore the two examples, are distinguished at least in so far as there is a degree of choice in the latter case which is absent from the former. It is not obvious that we are as free to choose which behavioural phenomena are criterial for toothache as doctors might be as regards defining angina. Again, the interchangeability feature seems much less readily applicable to the case of toothache. However, in both cases, once use is criterially fixed we no longer have any choice. Of course, we could speak differently but there is a clear sense in

which that would change the language we use, dependent upon the degree of difference involved. Therefore, both cases have the property of use-fixing in common and in this respect are identical. I return to the same point below, in a number of contexts.

The ninth use indicates and emphasises that doctors as practitioners of medical language-games may successfully and correctly use words in the complete absence of any such criteria/symptom distinction. Again the basic interchangeability of symptom and criterion is re-iterated:

"Doctors will use names of diseases without ever deciding which phenomena are to be taken as criteria and which as symptoms; and this need not be a deplorable lack of clarity. For remember that in general we don't use language according to strict rules-it hasn't been taught to us by means of strict rules either. We in our discussions on the other hand constantly compare language with a calculus proceeding according to strict rules. This is a very one-sided way of looking at language."
(Bl8 p.25)

In these final remarks it is made abundantly clear that in a certain sense criteria are irrelevant to correct use in this case and are by no means essential or necessarily pre-requisite. The ordinary business of language can readily proceed in the complete absence of criterial distinctions. At the heart of the very introduction of the terminology in question then is a claim which is antithetical to any suggestion that a theory of meaning was ever couched, explicitly or implicitly, in those terms. For here criteria just are not coextensive with meaningfulness. Criteria are not essential to meaning. Yet, to construe a notion as inessential, even frequently irrelevant, is plainly not equivalent to exposing that notion as illusory, bogus or inherently confused. This alone will not suffice

to establish the validity of Cook's interpretation.

Following Cook [1969] I now cite what was alleged by that author to be the conclusion to Wittgenstein's implicit argument:

"we are unable clearly to circumscribe the concepts
we use not because we don't know their real definition
but because there is no real definition to them."
(B1.B, p.25)

It is of vital importance to valid exegesis to examine the context of this remark. In fact it occurs in the third paragraph of p.25 which I have already quoted, above p.7. As noted that paragraph begins:

"This is a very one sided way of looking at language"
(B1B p.25)

But there is a definite continuation of sense, an enjambement, between the second and third paragraphs on page 25, as the initial of "This" makes obvious.

The opening "This" of the third paragraph transparently refers to the point immediately preceding, namely, that to demand that linguistic rules be calculus-like in their rigour and precision is an unrealistic and one-sided way of thinking about language. That claim arose when Wittgenstein emphasised the lack of necessity for criteria in every case and held correct use in the absence of any criteria/symptom distinction to be a linguistic commonplace. There Wittgenstein referred to a certain philosophical attitude which considered the lack of these distinctions as a lack of conceptual clarity. The point to which Cook refers may well be a rejoinder to that, rather Fregean, attitude and may well reinforce the point noted that many concepts do not have and need not have any explicit defining criteria. But that still does not show that no concept could have a criterion. It is just fallacious to infer from the fact that

many words do not have criteria to the conclusion that none could. As I read him Wittgenstein has just outlined what criteria for toothache or angina might be. What he is complaining about is surely that attitude which holds language to be a calculus and therefore demands precise distinctions in every instance. If anything, it is this attitude which is held to be inherently confused; not the concept of a criterion. Notably, Cook [1969] supplies his own reference for the "this" in question namely, 'common element', even although that phrase appears nowhere on either page. As Russell has asserted in another context "this may appear odd but that is not my fault".

Why one wonders, if Cook [1969] were correct, should Wittgenstein continue to invoke the term in later pages of the very text in which it is allegedly exploded? Undoubtedly, in the Blue Book Wittgenstein did think that as a matter of fact we possessed criteria for certain expressions. The concept is clearly used rather than mentioned. Comparison between different sets of criteria is frequently invited as in:

"(Here it will be useful if you consider the different criteria for what we call 'the identity of these objects'. (Bl.B p.55)

Again, describing the metaphysician's discontent with ordinary grammar Wittgenstein outlines the possibility of identifying cheating by means of criteria. Criteria are described as ways of finding out and are said to have a role in the meaning of these particular psychological concepts:

"The man who says 'only my pain is real' doesn't mean to say that he has found out by the common criteria, which give our words their common meanings, that the others who said they had pains were cheating. But what he rebels against is the use of this expression in connection with these criteria" (Bl.B p.57)

Further, why should Wittgenstein ever extend his usage of a 'bogus notion' beyond the Blue Book to later works ? If it were the case that the concept was destroyed in the Blue Book we should not expect any use of it, for example, in the Brown Book. The fact remains that nine uses are made of "criteria"/"criterion" in part one of the Brown Book. In each case one could fairly substitute, *salva veritate*, on that term with 'means of identification' or, more simply, 'way of telling'. For example:

"A criterion for what they mean would be the occasions on which the word we are inclined to translate as 'highest' is used, the role we might say which we observe this word to play in the life of the tribe."
(BB, p.94)

"All I can say is: If the fact that they only say "He can. .." if he has done ...is your criterion for the same meaning, then the two expressions have the same meaning."
(BB, p.104)

Equally, criteria are twice described as 'defining', as the criteria in the Blue Book were:

"When on the other hand we talk of the state of a stick in 46 observe that to this state there does not correspond a particular sense experience which lasts while the state lasts. Instead of that the defining criterion for something being in this state consists in certain tests."
(BB, p.101)

"As the defining criterion of "the chief picking men suitable for this action" I will not take what he says but only the other features of the situation"
(BB, p.104)

Interestingly, the notion features in rhetorical questions where the reader is clearly intended to conclude that what we do not have in the case in question is a criterion:

"'But doesn't B in this case lack the criterion by which he can recognize the material ? "
(BB, p.88)

"should we regard his personal experience as the criterion distinguishing between reading and not reading?"

(BB, p.122)

In each case the context makes clear that Wittgenstein is trying to show the reader the poverty of private mental states as public standards, as criteria. The reader would not be warranted in inferring from the fact that this or that is not, or could not, be a criterion to the conclusion that therefore nothing could be a criterion. Nor is it valid to argue from something's being a bogus criterion to the concept of criterion as itself a bogus one.

In part two of the Brown Book the term occurs four times during the discussion of 'seeing what is common': criteria for understanding are mentioned (p.132), as are criteria for 'seeing that' (p.135). A criterion for identifying two meanings of a word is suggested and behaviour is explicitly held to be criterial (p.138). Nowhere in this discussion of 'seeing what is common' are criteria ever identified as common elements, as essences or as 'what is in common'. Beyond that discussion, criteria as ways of identifying sincerity are mentioned (p.140) and the final use in the penultimate paragraph of the text asserts:

"If I speak of communicating a feeling to someone else, mustn't I in order to understand what I say know what I shall call the criterion of having succeeded in communicating?"

(BB, p.185)

Again, Typescript 229, published as Volume 1 of the "Remarks On The Philosophy of Psychology", contains eleven uses of the terms "criteria"/"criterion" (137 (2), 245, 302, 444, 530, 645, 649, 684, 698, 702). This typescript was dictated in Autumn 1947 and is a selection of remarks from a collection of manuscripts [M.S. books 130-138] which also formed the basis of Part II of the Investigations. According to the volume's editors G.E.M. Anscombe

and G.H. Von Wright:

"What Wittgenstein wrote in M.S. books 130-8 may with some justification be described as preparatory studies for Part II of the Investigations."
(preface)

Criteria feature in a number of different contexts in this text, as, for example, ways of telling whether someone can play chess:

"if he now replies that we are interested in what goes on in him after all, namely: in whether he can play chess-then we could contradict him by reminding him of the criteria which would prove his capacity to us."
(302)

Criteria are also presented as ways of telling whether a child has understood:

"...it is important that he ordinarily gives an answer when he has put up his hand; and that is the criterion for his understanding putting up one's hand"
(245)

Wittgenstein's first use in this text states that:

"I can't give criteria which put the presence of the sensation beyond doubt: that is to say there are no such criteria"
(137)

Of course, the fact that descriptions of the satisfaction of criteria do not entail the truth of assertions of the presence of the sensation does not itself entail that there are no criteria but the testimony that this remark provides cannot sit comfortably with the view that Wittgenstein thought of criteria purely as entailing, or as identical with sensation. I discuss this matter in more detail below. [Chapter 2, also Chapter 6].

There is a further degree of terminological similarity to the use of the term in the Blue Book in the last use in Typescript 229 (702) which mentions a "logical criterion". Further, in Typescript 232 [Sept/Oct 1984], published as Volume 2 of the "Remarks On The Philosophy of Psychology", criteria feature several times during discussions of aspect-perception. In

remark 42 Wittgenstein raises the question of whether there are criteria for perceiving changing aspects of a figure:

"if the child hit upon different ways of reproducing it [the figure] pictorially would that be our criterion for the change of aspect?"

In remark 43 Wittgenstein goes on to suggest that seeing might be a state but asks if this is so..

"..how do I know that a person is in such a state..what are the logical characteristics of such a state?"

The same thought is followed through in remark 44:

"For to say that one recognizes the state as a state whenever one is in it is nonsense. By what does one recognize it? (The criterion of identity)"

The criterion of identity mentioned in remark 44 would presumably be a logical characteristic of the state of seeing, as requested in remark 43. Such a criterion (if possible at all) would then also provide an answer to the question "how do I know.." in remark 43; in short a criterion would provide a way of telling. Remark 44 makes clear however that a criterion must be public and that, on pain of absurdity, no private phenomenon could function as a public criterion. In another remark on recognition in Typescript 229 (530) Wittgenstein makes the point that if all we have is a private experience of recognition without any public criterion for correct recognition then it may not make sense to speak of recognition here at all. Wittgenstein's use of "criterion" in remark 44, Typescript 232, has some of the tone of the Brown Book's rhetorical requests for criteria. The futility of any attempt to find a private criterion for public language might well explain the echo here of those earlier rhetorical uses.

The last use in the present typescript occurs in remark 506 which

concerns seeing aspects in geometry:

"..there are simply different criteria for seeing."

Nothing in any of the remarks composing the preparatory studies for Part I or Part II of the Investigations provides any evidence that Wittgenstein thought of the notion of a criterion as a bogus one. What clearly does emerge is a continuous thread of use of the term in the sense of, say, way of telling.

Finally, it must be asked why Cook himself continues to appeal to the concept of a criterion in his later article "The Metaphysics of Wittgenstein's On Certainty" [1985] if Wittgenstein had genuinely exposed the concept as confused. In that essay Cook [1985] readily relies upon the concept when, for example, he asserts: "...let us return to our criterion for picking out hinge-propositions" (p.90). Again on the question of the nature of hinge-propositions, Cook [1985] asks: "by what criterion does he [Wittgenstein] identify them?" (p.87). Lastly, Cook [1985] wonders of Wittgenstein had he: "ever worked out clearly in his own mind anything as definitive as a criterion?" (p.88).

Some explanation of why Cook does use the term "criterion" in that later article might be given on the basis of an argument in his earlier article "Human Beings" [1969], p.135:

"..if the idea of a 'defining criterion' is a bogus notion, then in considering the details of particular cases we are not considering merely 'symptoms'. These details can show us something about the grammar of the words...it is useful in philosophy to ask ourselves what..would enable us to recognize (would justify us in saying) that someone was expecting a visitor...it might be that the man was pacing up and down the room,...looking at his watch..talking about how good it will be to see his old friend again..if we were asked how we knew that this person expected a visitor, these are the sorts of things we would mention. The relevant point then, is that there is not some one thing, the same in all cases, that justifies our use of the word 'expecting', and if anything has the right to be called a criterion it is...such details of particular cases.

Presumably then, it is this second use of "criterion" as the details of particular cases which Cook [1985] continues to employ. Having said that, the final use I have quoted from Cook's [1985] article does concern "definitive" criteria and thus the explanation may not completely legitimate Cook's later use. However, in the passages quoted from "Human Beings" Cook's approach comes much closer to that of Baker [1974] and Hacker [1972] than his claim to have detected Wittgenstein's demolition of the concept in the Blue Book might lead us to expect.

In fact, it is the idea of a DEFINING CRITERION which Cook finds exploded in the Blue Book and not criteria of detail. Now, it is very plausible to distinguish between descriptions of the satisfaction of criteria which warrant a defeasible assertion of P and descriptions of defining characteristics of P which entail that P. Surely, one might then argue, if the characteristics are strictly definitional then they are not properly criterial. This is precisely Cook's [1969] move. I have no quarrel whatsoever with this distinction, which seems to me well-founded, or with the argument just as it stands, but the question remains as to whether Wittgenstein ever made such a distinction or presented such an argument. I have already shown how implausible it is to interpret the introductory passages as consisting of a refutation of the notion of a criterion as a defining characteristic understood in Cook's [1969] terms. The real basis for Cook's [1969] distinction centres on Wittgenstein's description of the criteria for angina, the second example, as "defining" while in the first example the criteria for toothache are not so described. In fact, what Wittgenstein says is that:

"'I have found the bacillus so-and-so in his blood' gives us the criterion, or what we may call the defining criterion of angina."

(B1.8, p.25)

But it is wholly imprudent to identify a new use on that basis alone. By parity of reasoning we must then recognise a further distinct use when Wittgenstein talks of the "logical criterion" [RPP Vol 1, (702) Zettel, (466) (477)]. Again, we are presumably dealing with a further use or two when Wittgenstein talks of the "real criterion" or the "genuine criterion" [PR, above p.1, para.1].

That is to say, in the introductory passages two examples of criteria are given. In each case Wittgenstein draws a distinction between criteria and "coincident phenomena" (ie. symptoms) and in each case the first use of the former term is simply "criteria" or "criterion". In the second case Wittgenstein does speak of "...the defining criterion of angina" but it must be incautious to interpret Wittgenstein's use of "defining criterion" as marking a new use in Cook's [1969] sense as equivalent to strictly definitional characteristics, if only because the latter is itself a highly technical term. The only formal term in the Blue Book passages, "tautology", is used very loosely in a non-technical sense. Similarly, I argue, "defining" is used in a loose sense and does not substantiate Cook's [1969] distinction. Unfortunately for Cook [1969], Wittgenstein does not explicitly condemn either example and appears to fail to make Cook's distinction at all. Therefore, it must follow that Wittgenstein just did not make the distinction which many commentators would have liked him to make and instead continued to use the concept of a criterion in a looser, more complex way which, for Wittgenstein, could legitimately include both the case of toothache and the more scientific case of angina. Certainly, the defeasible use has been perceived as the most interesting, and perhaps the most philosophically momentous, and that may well be so, but as regards WITTGENSTEIN'S use of the term criterion, as

regards what that concept meant to that author, let the use teach us the meaning.

III SOCIOLOGY.

I should say that I am aware of David Bloor's treatment of these passages in his: " Wittgenstein: A Social Theory Of Knowledge" [1983]. For the most part Bloor's [1983] account is faithfully explanatory. A criterion here is "an identifiable cue whose presence justifies the use of a word or classification" (p.41). This may be true but it might also be misleading. Because criteria justify use it might be thought that criterionless uses are therefore unjustified, incorrect etc. But that is just what Wittgenstein denies. Criteria are not presented as that in virtue of which we are justified in using words. We use words with no criteria without 'deplorable lack of clarity'. The point is: in general, to use words correctly we need no justification.

In this I think Bloor's [1983] account may be misleading but in another point it must be mistaken. Bloor [1983] argues:

"the empirical study of criteria would just be the empirical study of how words are given sense."
(p.46)

But if we can use words perfectly meaningfully and THEN establish defining criteria how can the criteria give the sense ? And in those cases where it is left open, where no criteria are defined, are we talking senselessly ? Clearly not. In discussing this point with me David Bloor defended himself from this criticism by distinguishing the viewpoint of the language-user from that of the spectator. This, I think, is fundamentally correct and is exemplified in the type of example with which Bloor [1983] is concerned (p.41- 42). The

standpoint of the social scientist, the anthropologist, is that of the spectator. The user's words are given sense, for the spectator, via the publicly manifest criterial relations. In that sense in which 'institution' is a sociological synonym of the philosopher's 'logic', criteria are part of the social institution of language. In the original case, if the spectator's study is a study of how words are given sense then it is a study only of how words are given sense for him. If this is so I do not think the original context makes it obvious. And, in any case, it does not exclude the point that for the language-user criteria just are not always necessary.

Finally, to ascertain whether there are criteria for a given word is to investigate the grammar of that word; not just to examine the phenomena empirically. It is a grammatical investigation not a purely empirical one. Therefore it is a philosophical rather than sociological investigation. If I know the relation between pain and pain behaviour is it something which I have observed? It must be at least that. But that is so if the relation is symptomatic. We learn criteria when we learn a language. To know of the connection empirically is not enough to know that the relation is criterial. Only an understanding of the grammar of a word, its use and role will tell us what the criteria for it are. Why this should be a matter exclusively for sociology and not for the Philosophy of Language I do not think Bloor [1983] makes clear enough.

IV USE.

What then is the significance of these introductory remarks? And what of the nature and status of a criterion for Wittgenstein rather

than his commentators ?

Wittgenstein's point in the remarks which introduce the concept of a criterion is, I argue, reminiscent of the famous point made by Bishop Berkeley [1980] against John Locke's [1975] doctrine of abstract ideas. As against the existence and conceivability of such ideas as potential mental images or icons Berkeley [1980] maintained that the generality in question, which Locke [1975] as semantic denotationalist saw as essential to explaining the generality inherent in language, did not consist in the nature of the ideas themselves but in their function or use. In practice, Berkeley [1980] argued, the abstract idea amounted to no more than a particular idea serving to illustrate certain general, common properties. Much as the geometer's particular line is used to illuminate the properties common to all lines. Here what is abstract about an idea consists not in its nature but in its use. Similarly, what is criterial about a criterion consists not in the nature of the phenomenon itself but rather in the use which a human community makes of it. What we use as a criterion is for us criterial. Ontologically, those phenomena which are used as criteria are ultimately actually empirically observable. But so are symptoms. It is the status of criteria in the institutional logic of language that gives them their special character and status. I have noted from the Blue Book that a symptom can become a criterion, that criteria can be established by convention and that criteria can change according to human purpose and practical convenience. Hence, I argue, Wittgenstein does recognize the existence of criteria and the legitimacy of such a concept but not in the sense of an essence or as strictly

definitional characteristics or indeed of anything which transcends human purpose and convenience. In the passages in question Wittgenstein is drawing attention to linguistic practices and is neither imposing upon nor deposing from any part of that practice. What is offered here is a descriptive picture of a fragment of language. If this is to be construed as a theory then it is a 'philosophical thesis' as delineated and commended in the Investigations and is as innocuous and unobjectionable as is there implied.

Simply and ordinarily, a criterion will be a useful and practical means of identification, of decision, which functions as a mode of everyday explanation. This characterization would seem very much in harmony with the widely acknowledged anthropocentric tone of much of Wittgenstein's later work. An equally simple account of criterion in the mouth of Wittgenstein is given by Colin McGinn [1984] as equivalent in sense to 'way of telling' but the suggested definition occurs in a footnote (p.124) and, unfortunately, no account of the Blue Book is given. Although I do not attribute my own account to that author there is no obvious incongruity between these views. Both will be found to fare equally and transparently well, I think, if the test of contextual intersubstitutivity is taken. The precise choice of words allows a degree of freedom but this, I firmly believe, is as much as Wittgenstein ever meant by "criterion".

V CONSISTENCY.

Having examined the use of 'criterion' in the Blue and Brown Books it is now possible to consider in a more detailed way the

source of Baker [1974] and Hacker's [1972] earlier claim that the Blue Book usage is in fact inconsistent.

Baker [1974] and Hacker's [1972] discomfort with the tone of the introductory passages in particular is comprehensible only in the light of their professed interest, at the time of writing, in developing a criterial semantics and their anxiety that descriptions of the satisfaction of criteria for P are represented here, in one case at least, as entailing that P. If descriptions of the satisfaction of criteria for P did entail that P then the criterial assertion could not be a defeasible one. To lose this would be to lose the most interesting and controversial feature of modern criterial theories. Taken together with the implication noted earlier that meaning and criteria clearly are not co-extensive in these passages a very dim picture emerges for any criterial semantics.

But criterial semantics is hardly Wittgenstein's concern here or in later works. Nowhere in the Blue Book does he assert or discuss defeasibility in his use of the term but neither does he elsewhere. The upshot is not that Wittgenstein had a very poor criterial semantics, as Carl Wellman [1981] has recently argued, but that he was never concerned to produce one and simply used the terms, "criteria"/ "criterion", in the innocuous sense I have outlined as, in McGinn's words, ways of telling. Hacker [1987] now recognizes this fact and refuses to take on board the idea of a Neo-Constructivist Theory of Meaning based on criteria:

"It was a mistake to present Wittgenstein's remarks about criteria as the foundations of a novel semantic theory, for he would have viewed the whole enterprise of constructing philosophical theories of meaning as yet another house of cards to be demolished not fostered."
(p.viii)

Despite this, some anxiety about consistency remains:

"In some of Wittgenstein's mathematical examples it is at any rate not obvious that what he calls a criterion falls short of entailment..Similar considerations apply to the angina example."
(p.316)

The extent of the revision in Hacker's [1987] revised edition is considerable, as is the degree of improvement in the essay on criteria therein. A brief account of the new conception will make clear why the the claim to inconsistency remains; albeit relegated to a footnote. In the final chapter of the revised text Hacker [1987] attempts to find a middle path between two accounts of criteria. Firstly, Hacker [1987] rejects McDowell's [1982] criteria which exemplify pain directly, such that no "logical space" mediates between the criteria and the pain. Hacker [1987] contends that this denies any inferential structure to the relation and is a costly equivocation as a strong form of behaviourism is an obvious logical consequence. However this may well be more Hacker's characterisation of McDowell [1982] than actually McDowell's [1982] view. For although McDowell [1982] does assert an identification, in a sense, (a sense more fully discussed below, Ch.6) it is not an identification involving pure behaviour but rather PAIN-behaviour.

Of course, there is only behaviour, but not colourless bodily movement, for the description of the behaviour ineradicably involves mental vocabulary. Therefore, no behaviouristic reduction can be effected. The real trouble here, which Hacker [1987] does not make clear, is that instead of having to make a case for criterial support ever adding up to truth the problem becomes that of explaining how, on the basis of criteria, we can ever be wrong.

On the other hand Hacker [1987] wants to distance himself from

Shoemaker's [1963] conception of logically necessarily good evidence. This is perhaps the classic, post Wittgensteinian, characterisation of criteria which provides the terms of art for Gregory Lycan's [1974] scholarly synopsis of the field. Hacker's [1987] worry here is that this might make the distance between the relata too great. Criterial inference seems to Hacker to become inference from the physical to the spectral. Again, I argue, the real difficulty is not made clear. What Hacker [1987] does not mention is that if pain behaviour is logically necessarily good evidence for pain then there is no possible world in which it is not evidence for pain. To make use of these notions is really to commit oneself to saying something about how they are being used. For example, the concept of remoteness requires further explanation.

In Hacker's own revised view a criterion for P is best described as a:

"grammatically determined ground or reason for the truth of P."
(p.315)

For Hacker [1987], descriptions of the satisfaction of P's criteria never entail P. One may feign pain. Beginner's luck can be mistaken for genuine ability. It is because criteria are circumstance-dependent, Hacker [1987] argues, that criterial assertions are essentially defeasible. These considerations raise deep problems all of which I examine in detail below [ie. Ch. 5, 6 & 7]. Despite this it is worth pausing briefly to consider what Hacker means by 'defeasible'. In the literature this has almost become a technical term in the explanation of the concept of a criterion and is rarely itself explained. I wish only to note the complexity of the notion. Firstly, as regards pretence, although the criterial assertion of

pain is defeated it is not the case that the criteria for pain are destroyed. Pretence gets off the ground on the back of pain-criteria, by exploiting rather than destroying those criteria. Again, there may be good intuitive grounds for a distinction between defeating evidence and overriding evidence. For certain purposes there may be no need to emphasise this complexity but this need not necessarily be so in every context.

The general point which Hacker is making here is the logical one that the criterial relation is Non-Monotonic. That is, it is such that when more information is conjoined to the premisses, the conclusion may no longer be warranted. For example, on witnessing pain behaviour one judges that this is a case of pain. If one then discovers that this is in fact a case, say, of acting, or of grief one is no longer entitled to infer to pain.

Finally, Hacker notes that neither what is to count as a criterion nor what is to count as a defeater can be completely described in advance of particular contexts. We can only teach by reference to paradigms and thereby enable deviations from the norm to be recognised. Again this is a logical consequence of the circumstance-dependance noted earlier. These points in particular are discussed below in more detail but for present purposes this much is all that is required. That is, it can be clearly seen that it is of the essence of the account of criteria which Hacker [1987] attributes to Wittgenstein that criterial assertions be defeasible assertions. Hacker is very clear about this in both editions of "Insight and Illusion"; much clearer in fact than Wittgenstein himself. It is this essential feature, I argue, which Hacker feels is threatened by

the example of angina in the Blue Book.

In the introductory remarks to which I have referred discussion centres on two examples, namely, that of toothache and that of angina. In each case Wittgenstein speaks of criteria in connection with the meaning of the term. Many commentators (ie. Albritton [1970], Baker [1974], Hacker [1972]) have held that there is an important difference between the two cases. That difference, as I have noted, concerns the issue of whether or not a description of the satisfaction of criteria for P actually entails P. The criteria themselves, of course, entail nothing for the trivial reason that they are non-linguistic. The other side of the same coin is of course whether or not an assertion based on criteria is defeasible. If the assertion is defeasible it follows that it should be quite possible for the criteria to be satisfied without that for which they are criterial being in fact the case. If the description of the satisfaction of P's criteria logically entails that P this is just what is not possible. The problem, and the root of the inconsistency allegation, lies in the fact that while clutching one's cheek and having a red patch upon it is consistent with not having toothache at all a description of the presence of a bacillus whose presence has been taken to define angina seems to imply that one has angina. From this premise it has been inferred that Wittgenstein is inconsistent about criteria (ie. Baker [1974], Hacker [1972]) or, more charitably, that he has at least two distinct and incompatible notions of a criterion (ie. Albritton [1970]).

But the matter is not quite as simple as this. Firstly, it is not immediately obvious to me that the angina case is one of

entailment. It surely makes perfect sense to say that it may be that the bacillus is taken to be defining and yet there might be empirical pressure to alter our present definitions, say, some people who have the bacillus do not have angina or again that there are cases of angina in which the bacillus is not present. If this were the case doctors would clearly revise the original definition. From a contemporary standpoint it can be argued that the empirical pressure is to change definitions but also that defining characteristics cannot be criteria on pain of sacrificing the defeasible nature of criterial assertion. Hence authors such as Cook [1969] might conclude that the angina case is not a criterial one. As I have noted Wittgenstein does not explicitly discuss this point. But that need not imply that his use is inconsistent. That would be so only if the term "defining criterion" is used in the precise technical sense of defining characteristics. I have already warned against such an interpretation. It is by no means clear that Wittgenstein's discussion is this precise or technical and it does not help simply to assume that it is.

Even if it can be taken for granted that the second case is entailing while the first is not it remains invalid, I argue, to infer that Wittgenstein's concept of a criterion must be inconsistent. Both recognising that a defining characteristic obtains and recognising the satisfaction of criteria for P motivate assertions that P. Both are 'ways of telling', 'cues for use', means of identification. What has come to be defined as a criterion is 'non-inductive evidence' (ie. Lycan [1974]) which grounds defeasible assertions. But if the second case is entailing then 'criterion' in Wittgenstein's later work may mark either type of case. If we are

being presented with two different ways of telling, means of identification or cues for use it remains the case that both are ways of telling. The notion may just be complex enough to allow different things to count as ways of telling, as criteria. The degree of precision required of a way of telling would, for Wittgenstein, be wholly purpose-and context-dependant. It is no surprise then that the exact character of the method involved will alter according to just those factors as appropriate. The insistence of some commentators on there being an inconsistency arises, I argue, when they are looking for a Theory, an essence. If, focussing on one case, we take it to be essentially criterial, as contemporary orthodoxy precisely does, then any deviation from that case will look inconsistent with it. This is exactly what Baker [1974] and Hacker [1972] have done. Having good reason to prefer a defeasible model that is what they extracted. The rest is condemned as inconsistent. It is a platitude to warn against foisting a frame on any of Wittgenstein's work. If we are to see what was meant let us look for the use; the range of the remarks. It is no part of the investigation that it will culminate in a Theory. Examining Wittgenstein's use of "language-game" will no more tell us the essence of a language game than examining his use of "criterion" will illumine his Theory of Criteria.

Anxiety about the strength of the relation as presented in the Blue Book is common to a number of the most influential commentators and is again evident in R. Albritton's [1970] article on criteria which may miscast the relation. Here Albritton [1970] outlines two uses of the term criterion in Wittgenstein's later writings. Firstly,

Albritton's Wittgenstein asserts that:

"the criterion is a logically necessary as well as a logically sufficient condition of its [ie. that for which the criterion provides support] being so.."
(p.235)

Secondly, the stronger claim is made that:

"..to be a criterion of X is just to be what is called X"
(p.241)

In both accounts then Albritton [1970] asserts a much stronger relation than Hacker [1972] might have wished for. Clearly, a description of the satisfaction of P's criterial circumstances simply entails that P on this model. But is this what Wittgenstein had in mind here ? It is illuminating to apply some of Wittgenstein's own methodology here and to focus attention upon the philosophical grammar, the form of language in which Albritton's [1970] account is couched. In both accounts Albritton [1970] plays a very formal, technical game employing the language of entailment and of necessity and sufficiency which could not be more stylistically distant from Wittgenstein's own mode of expression. In fact, in the central text which Albritton acknowledges (Bl.8, p.24-25) the terms 'entail', 'necessity' and 'sufficiency' do not occur at all. On this point Albritton [1970] complains that:

"..if 'logically necessary and sufficient condition' and 'entails' are jargon, they have the company in that misery of the word 'tautology', which Wittgenstein uses in the passage I have been discussing."
(p.236)

But this does not alter the fact that the whole tone of those passages is distinctly non-technical and devoid of any of Albritton's [1970] jargon. Indeed, the term 'tautology' does occur. But that is the only technical term which occurs and even this is used in a non-technical sense or, at least, is not obviously used in a technical sense. Again, I have suggested that phrase 'defining criterion' is

also used loosely and may not be straightforwardly equivalent to "defining characteristics". Instead, Wittgenstein explains a criterion more like a rule for use, as an aspect of the grammar, use or meaning of the term X which rests, fundamentally, on convention. In employing the language of formal logic, I argue, Albritton [1970] gives a misleading picture of what Wittgenstein understood a criterion to be. I have already noted that defeasibility might be involved in one of the two introductory examples and cast doubt on the appropriateness of the language of entailment to the second. The context in which the introductory discussion of the term emerges is the Blue Book which proceeds from the question 'What is the meaning of a word ?' The discussion is embedded in an attempt to describe the ordinary grammar of language.

In his first 'use' Albritton [1970] shows a tendency to logicise and over-emphasise aspects of Wittgenstein's use of the term and I think falls prey to just that "talk of language as a symbolism in an exact calculus" (Bl.B, p.25) of which Wittgenstein warns: "the puzzles which we try to remove always spring from just this attitude to language" (Bl.B, p.26). The second of Albritton's [1970] uses which asserts an identification between a satisfied criterion of X and being X itself is even more misleading. What it does make clear is one way in which commentators may have come to the view that Wittgenstein was a behaviourist. For, if we substitute Albritton's [1970] account in psychological contexts, his identification might well generate a very strong form of behaviourism. That is, if the criterial basis for, say, pain-ascriptions, is held to be exhaustively describable in purely behavioural vocabulary. Otherwise, of course, no reduction

would be effected. The refutation of the suggestion that Wittgenstein was a behaviourist and a full critique of this type of commentary requires an examination of Wittgenstein's later philosophy of mind. The following chapter therefore proceeds from a paradox at the heart of that philosophy of mind to a discussion in the third chapter of the role of criteria in the Investigations.

CHAPTER TWO

INTRODUCTION: A PARADOX IN WITTGENSTEIN'S LATER PHILOSOPHY OF MIND

The paradox I draw attention to in the chapter below is not the one made famous (or infamous) by Saul Kripke in Chapter Two of his "Wittgenstein on Rules and Private language" [1982]. Kripke's [1982] "Wittgensteinian Paradox" is a fundamentally sceptical paradox allegedly upheld in remark 201 of the Philosophical Investigations. Although there is an apparent retraction in footnote 7 [Ch. 1]:

"That discussion may contain more of Kripke's argumentation in support of Wittgenstein rather than exposition of Wittgenstein's own argument.." (p.6)

the retraction only concerns Kripke's account of "the dispositional theory". For Kripke [1982] the pedigree of the sceptical paradox remains unquestionably "Wittgensteinian":

"I urge the reader to concentrate...on...the intuitive force of Wittgenstein's sceptical problem."

But the sceptical problem is not Wittgenstein's. In the second paragraph of 201, which Kripke [1982] does not cite, Wittgenstein explicitly says of the sceptical paradox:

"But it can be seen that there is a misunderstanding here.."

This description of scepticism in general as the product of misunderstanding is far more characteristic of the writings of that philosopher at every stage than is support for any 'sceptical paradox'. I discuss this issue more fully below [Ch. 3, Section II]. But in any case the fact remains that the paradox is not being upheld by Wittgenstein in remark 201. The issues raised by Kripke do however merit closer attention despite being more Kripkensteinian

than Wittgensteinian and are discussed in Chapter 5 below. I do not therefore discuss the 'Wittgensteinian Paradox' at this stage but rather Wittgenstein's paradox which is baldly stated in remark 304 where Wittgenstein asserts of a sensation:

"it is not a something but not a nothing either."

That remark also contains the assurance that the paradox can be dissolved; a claim which represents an important challenge from the standpoint of Wittgensteinian interpretation.

I THE LANGUAGE OF SENSATION

At first blush the most notable feature of remark 304 is the relative scarcity of discussion of this paradox in the literature as compared to Kripke's paradox [P.I., 201]. This state of affairs is not in itself disadvantageous to the present inquiry. In fact it may be of some methodological advantage in permitting more exclusive concern with the remarks of Wittgenstein rather than his commentators. The key to the correct interpretation of the first half of the paradox, I argue, lies in remark 293. There Wittgenstein is concerned to argue against a particular conception of the language of sensation which describes the relation between sensation-terms and the sensations themselves on the model of names to their bearers and, more specifically, one which construes reference i.e. the naming relation, as fundamental to semantics. On that view a sensation term is a label on a private object. Only by owning the object can one have any idea of the meaning of the label. Further, because the object is private the label can be fixed only by the object's owner. Hence, label-fixing must be a matter of

private ostensive definition.

Given Wittgenstein's attitude to Hume, as recorded by Norman Malcolm in his "Memoir" [1958], he may not have had "A Treatise of Human Nature" [1985] in mind here. Yet, I think, Hume's doctrine of impressions and ideas represents the type of position which Wittgenstein did have in mind. Hume speaks freely of impressions of sensation as internal objects which cause mirror-like ideas. The idea in turn ostensively defines a word which acts as the sign of the idea. The choice of a particular sign is ultimately arbitrary (ie. it is not a "natural" sign). Once established it is, of course, held constant. Without the idea as internal object we can literally have no idea of the meaning of the sensation-term. In order to create the idea in another we generate the impression, repeat the name and thereby promote an introspective association; a private ostensive definition. The doctrine of impressions and ideas provides the positivistic rationale for Hume's approach to causation. To verify necessary connection in experience would be to have an impression which would give rise to an idea and in this way some actual content would have been given to the expression "necessary connection". The doctrine of impressions and ideas therefore brings together Hume's epistemology, semantic theory and philosophical psychology.

Again, the theory of language and classic empiricist psychology of John Locke's "Enquiry Concerning Human Understanding" [1975] exemplifies the type of philosophical thinking which Wittgenstein wanted to challenge. The first axiom of Locke's theory of language is that words stand for ideas and therefore sensation terms stand

for ideas of sensation. Commentators such as Jenkins [1983] have argued that words have a dual role in Locke's philosophy of language, since they also have a use in referring to objects. In the psychological case however this will collapse into the single purpose of referring to ideas of sensation. These ideas reflect the impressions of sensation which are their causes and in virtue of the representational principle at work in Locke generally we have a very clear example of the natural signification between the idea as meaning and the impression which is its cause.

Wittgenstein's argument in remark 293 of the Investigations points to the conclusion that if the game is played at all, if our practice with and use of sensation-language can go on at all, then the whole story about private objects becomes superfluous. Integral to the argument is the claim that if we conceive of the situation as the classic empiricist invites us to then there can be no guarantee either that the private object will not vary in character from subject to subject, or that the impression might not be different in every case. It seems equally conceivable on the empiricist model that the object might even constantly change in every case for an individual subject. On that view what we lack is any criterion for sameness and again any criterion for correctness. Most importantly, we just have no criteria whatsoever on this model. As Bernard Gert argues in his "Wittgenstein's Private Language Arguments" [1986, p.419]:

"some..think that a private language is impossible because there can be no independant check on my memory. But Wittgenstein's point goes deeper than that; it is not that "I have no criteria of correctness" but that "there are no criteria of correctness". This is the reason that "whatever is going to seem right to me is right". Because there is nothing to remember that will tell me

how the sound is to be used."

Gert's comment is not confined to remark 293 but, as the title suggests, ranges over the full breadth of "...the private language arguments" which are taken to include 293. In that context Gert's point is that there is nothing about the sensation as private object which yields any clue to the use of the sensation-term. Again, it is not just that introspection provides a poor criterion, or that any particular criterion is lacking in that context. It is rather that no criterion whatsoever has been established. If we hold these considerations in abeyance, momentarily, and generalize universally on the naming-theory model, over the domain of the set of sensation-terms, we arrive at a picture of a linguistic community who only know and understand the language of sensation by reference to their own sensations and can never ultimately know when or whether fellow language-users ever entertain such experiences. Ultimately, a form of scepticism about other minds is generated for as Wittgenstein argues: "the box might even be empty." (P.I. 293).

The form of the argument in this remark is characteristic almost to the point of definition of Wittgensteinian criticism in being a *Reductio Ad Absurdum*, and in this instance it is clearly a sound example of such. Sensation-terms cannot simply be substantives whose meanings are provided by their bearers precisely because the independent practice of sensation talk could proceed if such bearers constantly altered or even disappeared altogether. The point here is not even that there cannot be any genuine names in the language of sensation but rather that these names cannot acquire their role and status in the way in which the refuted view implies. We can never justifiably jump from the existence of names directly

to their actual use which must be taught and learned. This I hope is something like Wittgenstein's reasoning here.

The argument cuts across both the Empiricist psychology of Hume and Locke and applies equally to Cartesian Rationalism. In general, to so construe the language of sensation is to condemn the very sensations to irrelevance. The fundamental error resides in the misguided application of the grammar of one language-game, physical-object designation, to a wholly distinct language-game, that of sensation. Construed in that way the language of sensation presupposes objects of reference but as no public, physical object exists private mental objects must be postulated. It is, I believe, in this sense that sensations are not somethings. That is, however the contents of consciousness are to be portrayed they must not be pictured as quasi-substantial objects floating in an ethereal realm awaiting a Humean nomenclator. As I have argued, postulating such objects is logically insufficient to account for the linguistic practices we possess. Neither can the privacy survive the critique. If it is possible to identify a peculiar feeling in oneself, say, as might be experienced in a lift due to shifting in one's centre of gravity, and to give the peculiar feeling a name, then that possibility itself already involves and presupposes a complex background of conceptual apparatus: understanding the concept of a lift, of motion, of sensation, knowledge of what 'naming' is etc. The price to be paid if that is denied is again that we are left with no criteria whatsoever. In short, the privacy of the object can never explain the public use of the language.

These conclusions need not imply that sensation language is

never descriptive. The relation between the sensation-language and the sensation might perfectly well be held to be genuinely descriptive but very different things can be subsumed under that heading. A description of a room and a description of a sensation can both be legitimately held to be descriptions but they will also be seen to differ importantly. I can, for example, be wrong about a feature of the room I describe but it is not clear that I can be wrong about a description of my own sensation. Could I be wrong about my being in pain ? Suppose a malignant psychologist creates an environment which appears to contain a red-hot stove. In fact the stove is cold. My hand is forced down upon it and I scream out that I am in pain. Even here it is not clear that I am wrong. It was undoubtedly an unpleasant sensation, perhaps a Humean sympathetic pain. Again, it makes clear sense to doubt the accuracy of my description of a room but does it make sense to be in pain and to doubt or question that fact ? If there is no room for error and no role for doubt then is there room for correctness or a role for knowledge ? Such argument is thematic of the Investigations and is, I argue, directly responsible for Wittgenstein's remark in 246 that to have said of me that I know that I am in pain is to have said nothing more than that I am in pain. The same point is reiterated in On Certainty in remark 504:

"Whether I know something depends on whether the evidence backs me up or contradicts me. For to say one knows one has a pain means nothing."

In one sense the claim is a simple one about ordinary usage. As opposed to ordinary avowals "I'm in pain !", "that hurts !" etc the claim " I know I'm in pain" is, at the very least, an unusual one. In fact, it is more likely to be heard in departments of

philosophy as an utterance from someone in debate rather than pain. The fact that, with sufficient artfulness, special contexts can be imagined (such as the philosophical one) which do confer sense upon the utterance does not refute the simple point about ordinary contexts. The remark from *On Certainty* provides a little more background here however and is indicative of a more general theme about the absence of contrasts, the absence of a role for evidence, of any way of coming to know and is intended, I argue, to cast doubt upon the content of knowledge-claims under such circumstances. Wittgenstein's point might equally well have been put, I suggest, by the claim that it is because we can neither know nor not know that to seriously talk of someone knowing of their own sensations is, in an important sense, redundant. But the claim in question is not as revisionary as might be thought. It involves an acknowledgement of the authority, the special status of the individual, as regards his own sensations but couples this point with the claim that unless there are public criteria in terms of which we learn the language of sensation in the first place then a radical scepticism follows, as the story about private objects remains logically insufficient to explain our use of a public language. However, while it is in terms of criteria that we learn the language of sensation, once that is learned we do not identify our own sensations by observing that we satisfy the criteria. This places the authority of the individual in his own case and the first person/third person asymmetry against the backdrop of a logically primary common public language. The private is therefore dependant upon the public in a sense which reverses the logical priorities of traditional Cartesian

epistemology. Despite this, Wittgenstein nonetheless seems to allow that while doubt has a role in the third person case it has no role in the first person case. In that he might well agree with Descartes except in so far as he takes the dupitable to be logically and epistemologically prior to the indupitable. But if doubt is to be excluded from the first person case because it has no role then so, Wittgenstein seems to suggest, must its counterpart knowledge. I critically assess and discuss these points in more detail in Chapters 6 and 7 below but this, I argue, explains Wittgenstein's prohibition against knowledge-claims in first person cases.

As a final point against the type of view considered in remark 293 it may be argued that the inner-sense cognitivism which Modern Philosophy inherits, ultimately, from Descartes and which is clearly illustrated in classic British Empiricism misconstrues the nature of personal experience. For there is no cognition of being in pain via inner sense; there is just the state of being in pain. That state already involves the awareness; pain is itself a form of awareness. To postulate a second order awareness from inner-sense is to run the risk of a third order, awareness of that awareness and so on. The implication of my account of Wittgenstein is that it is in the first instance on the basis of criteria determining the meanings of sensation terms that we learn to discriminate between, classify, name and describe our awareness.

While I would not want to argue that all our sensations are invariably conscious sensations (it seems clear to me that many of them are not) I would want to argue that it is always possible that a sensation be conscious. That is, I do not want to rule out my automatically and unconsciously brushing an insect off my hand when

reading, for example, though I would want to say that that sensation could be a conscious one. Therefore, I would want to argue that a sensation which was not possibly consciously manifestible is no sensation at all. To be in pain but to be unaware of the pain might be a poor description of being anaesthetised but it would not be a description of being in pain. D. Dennett [1981] has recently argued that such a state of being in pain but lacking any awareness of the pain could describe the result of performing a lobotomy on a subject. Norman Malcolm has held this claim to be a bizarre one. If I subtract the felt character of the pain from the pain can I be confident that what I have left over is still pain? Malcolm's doubt is a reasonable one supported further by the fact that neither Dennett nor those who authorize or perform such horrendous operations can possibly know that the result of it is truly described as has been claimed, pain without the awareness; if, indeed, any sense can be made of the description at all. The operation is performed in the last resort to modify a subject's behavioural reactions. It may be a more pessimistic view that the subject still experiences horrendous pain and distress although he is no longer physically able to express it but it is not one which Dennett, or logic alone, can simply rule out. That is to say, while I would want to allow that there must be some outward manifestation, in principle at least, it remains possible that the effect of the lobotomy is on a par with a paralysing poison, a sophisticated, restraining rope.

Neither, I argue, will the original description be truly applicable to animals. Certainly, cats lack conceptual thought and

the ability to express pain linguistically. The fact that the cat cannot say: "I know I'm in pain", if that is taken as a marker of conceptual thought, will not lessen for a moment the excruciatingly painful character of the sensations it feels in a laboratory experiment, or in having a leg or tail broken by a motorist. Any lack of awareness here is due not to the sentient animal but to those human beings who are too insensitive to recognize the suffering, or to allow it to count. Finally, of those human beings who cannot feel pain we do not say that they are in pain but do not know that they are for it is just the being in pain which they lack. They simply do not have the sensation, the awareness. As Geach [1967] argues, what we are aware of is the pain not the cognition. The analogy to sense suggests the possibility of error. But I think serious doubt has already been cast upon that suggestion.

In concluding my treatment of the first half of the paradox it might be felt that the case against epistemic and semantic privacy is not fully prosecuted just on the basis of the argument in remark 293. To this I concede that a fuller appreciation of the suaviseness of Wittgenstein's critique of meaning as mental process requires a much wider survey of his remarks and it is to this that I now turn.

II Experience And Meaning

Remark 293 is instructive in a fundamentally negativistic sense. It considers and refutes one answer to the question: what is it for a word to mean? At best, it illuminates what sensations are not by pointing out one meaning model which cannot be properly applied to the language of sensation. The *Reductio Ad Absurdum* is a very

characteristic form of Wittgensteinean criticism and as a mode of instruction has its purpose. In order to clarify the later philosophy of mind so as to refute the behaviourist allegation I continue to refer to it after absorbing the question above into the inquiry as to what it is for a person to understand a word, much as Wittgenstein himself does in the Blue Book. I examine the next move to meaning as use and the question as to what it is to use, to follow a rule, in Chapter 5 below. At this stage I want to discuss the idea of a 'private language argument'. This will bring sharply into focus possible answers to the question as to what it is for someone to understand a word. The range and critical power of Wittgenstein's remarks on that subject will refute any allegation that the targets of those remarks are unsophisticated.

The critical target of remark 293 is the basically Augustinean position which asserts that language is essentially a matter of the composition of names and explains meaning in terms of that naming-relation. Concern with such positions is thematic of the Investigations. Augustine, and the dream theory outlined by Socrates in Plato's Theaetetus, are explicitly quoted; and the Atomistic views of Russell on the logically proper name, for example, are examined at some length. The first 120 remarks constitute strong evidence for this interpretation. But with 693 remarks in part one alone whether that view can be extended as an explanatory device of value throughout remains an open question. By indicating where and how more sophisticated positions are rebutted I argue against that view. What follows cannot be a wholly comprehensive exegesis but should be sufficient to establish the point.

The account of sensation-language which I considered in remark 293 identifies the sense of a sensation-term with the sensation as quasi-substantial object and therefore makes meaning a property first and foremost of the individual rather than the community. One might try to preserve this latter feature, this privacy of meaning, and improve upon the plausibility of the account of the meaning of sensation-terms by identifying the sense of a sensation-term not with the object which it allegedly names but with the idea which, for example, the empiricists took to be caused by the experience of the object. On such a view sense will still be private and for the individual will consist in, let us say, mental representations; construed very broadly. These turbulences in the stream of consciousness could be construed as inscrutable to other language-users and in this way privacy of sense could be ensured. In order to guarantee the possibility, indeed the empirical reality, of communication a coincidence or convergence of aspects of sense might be postulated as a bridging-mechanism from private sense to public meaning.

This position is clearly more sophisticated than its predecessor and yet Wittgenstein does deploy a range of arguments against just such a view. The first of these draws attention to grammatical differences between understanding and mental processes of any kind and is an important wedge-driving preliminary to a sustained critique. At p. 59 Wittgenstein indicates that clear cases of mental processes differ categorially from those of understanding both in terms of intensity-aspects and the nature of their temporality. Understanding, Wittgenstein argues, is not a process term but is more akin to an ability or disposition. It is a technique

exemplified in behaviour rather than a psychologically continuous experience. The point is not that understanding completely lacks experiential content but rather that as mastery of a technique it involves more than that.

A second group of arguments offers further support for the claim that what comes before the mind cannot constitute a sufficient condition of meaning or understanding (139, 141, 151-2). Here it is emphasised that understanding the meaning of a word must include a working knowledge of the roles played by it in the network of appropriate contexts and indicates the inadequacy of mental processes to the task. Sense could not consist in an image, feeling or formula for possession of that will not of itself teach the application of the word; it will tell nothing about what to do next with the word. Even if a given picture of, say, a cube suggests a use, ie. its applicability to cubes, the individual is still free to employ a method of projection under which that image applies to triangular prisms. A line sketch of an old man walking up a hill with a stick is just as validly a line sketch of an old man sliding backwards down a hill. The point being that any compulsion from an image to a specific word-use is psychological and not logical. We possess two distinct sets of criteria: one for representations, the other for their application. The picture need not even be imaginary. For if the use is known the picture's task could as well be fulfilled by a public drawing or physical model. Again, although the same image occurs on two occasions of use neither the word nor the meaning of the word need be identical in both.

A third argument-group contests the claim that any mental process

constitutes a necessary condition of understanding (151-2, 329-30). For example, a thought-experiment is outlined that involves firstly, both saying and meaning a word and secondly, simply privately meaning it. The point is that there is nothing left of the meaning of the sentence once we subtract the use of the sentence. The experiment is designed to make clear that it is just not true that mental processes invariably accompany verbal expressions: meaning is not a process and, a fortiori, no process invariably accompanies meaning. Again, Wittgenstein is denying the existence of a second-order, semantic language existing only in the mental realm. That is again his point at p.18e when he asks:

"Can I say 'bububu' and mean 'if it doesn't rain I shall go for a walk'?"

The same conclusion is endorsed in remark 329:

"language is the vehicle of thought."

This much classification of Wittgenstein's remarks is available in McGinn [1984] though I do not wish to blame that author for the account presented here. While I follow McGinn in not laying claim to a comprehensive classification, I argue that it must be broadened at least as far as a fourth category. This fourth group of arguments is best approached from a question already raised by remark 293: how do sensations or mental processes ever become associated with words either as the referents of names or as the senses of expressions? It might be claimed that by some reflective mental act of deliberative linking a sensation is named or a process associated with an expression. But is this even possible? Can the terms 'named', 'defined', 'associated' be understood at all in a context where the necessary backdrop to those practices is wholly absent?

If, for instance, the mental act is supposed to be one of definition then unless it determines future application it cannot be made sense of as definition. But this is just what the mental act cannot do. Again, following Gert [1986], it is not that "I have no criteria of correctness" but that "there are no criteria of correctness". There is no external court of appeal for correctness of use and:

"Whatever is going to seem right is right.."
(P.I. 258)

Whatever the act amounts to it cannot amount to definition. Worse still, similar consequences will arise for 'naming' and 'association'. In order to explain the use of any of these terms we must presuppose exactly what we seek to explain, namely, an established use. Again, while an act of this kind might be attempted and might even be sincerely felt a success it remains possible that in fact it has not been properly understood, possible that subsequent actual use will exemplify misunderstanding. Here public use corrects private opinion for meaning can only be taught by use in context. It is also reasonable to expect that if sense is a mental occurrence then that occurrence should at least be co-extensive with meaningful utterance. But even if this implausible claim were actually demonstrable the claim that it is precisely in this that understanding consists would surely require independent argumentative support. It is natural to wonder why we need postulate any coincidence or convergence of sense between language-users as guarantor of public meaning when rules of use and application are evidently common anyway. No real content has been given to the notion of a convergence which looks like a prime candidate for the application of Occam's razor.

Meaning cannot therefore be identified with mental processes but despite this, as Lewis [1980] notes, in Part II of the Philosophical Investigations [ie. p.175, p.241] Wittgenstein suggests that the mental accompaniments of utterances have a vital role in giving life to those utterances, personalizing use and giving substance to idiolect, to dialect. Although they do not exhaust meaning because as rules are shared public communication could continue. There is strong reason to believe that although language could continue as "basic communication", in Strawson's terms, linguistic and perceptual capabilities might be seriously restricted. The breadth and quality of our perceptions would be reduced by the semantic equivalent of aspect-blindness. To so lack the mental accompaniments of a language is described by Wittgenstein in Zettel (183-4) as "meaning-blindness". Characteristically, the meaning-blind would be unable, for example, to understand ambiguous words in contexts which were not clearly disambiguated, would be unable to mean the same word in different senses, to comprehend secondary senses of terms and therefore would be incapable of appreciating the sadness of music. The meaning-blind would be a:

"..people without stories, poetry, imagery, songs.."

and therefore, as Lewis [1980] indicates, would be a radically different group of language-users. The point is that though communication would be possible for the meaning-blind it could never amount to more than mere communication. Wittgenstein is not therefore fairly described as reducing the private, the mental to the public. Nor, I shall argue in the subsequent section, is he guilty of reducing the mental to the physical. Personal experience

is vital to language and meaning even though in the final analysis communication via public shared meaning, rules of use, underwrites the existence of language.

My classification is not by any means an exhaustive one. But it is sufficient to warrant the rejection of the claim that the Investigations deals only with the obsolete and the untenable. It is also sufficient to establish a further point or two about that work.

Many of the remarks which make up the Investigations exemplify a striking unity of subject matter and this is illustrated clearly in remark 293. In semantic terms, it argues that the language of sensation cannot be underpinned by private meaning-bodies as the bearers of names, while in mentalistic terms, sensations themselves are no sort of private object, no sort of object at all in fact. Though this remark and many others are readily described as philosophical psychology they are as easily subsumed under the headings of language or meaning. The author himself speaks here of continually converging upon the same problems from different perspectives:

"The same or almost the same points were always
being approached from different directions..
(P.I., preface)

The particularity and apparently disparate character of the remarks together with their emphasis on specifics of context mask that feature and can be misleading.

Further, the fact that remark 293 and all the others cited argue against kinds of private language without every remark being THE private language argument must endorse the conclusion that none of them is. Precisely because the critique of privacy of meaning,

privacy of language, is not ultimately any single argument but is a thematic concern which runs throughout the Investigations. Philosophical energies might be more profitably employed identifying and distinguishing the targets of these remarks rather than awarding them titles.

III BEHAVIOUR AND SENSATION

According to my account of the first half of the paradox Wittgenstein denied that sensations could be properly understood as private mental objects but in so doing did not deny any role for the mental accompaniments of language-use. Neither, I argue, did he identify or reduce the mental to the physical as regards behaviour and sensation. However, Rogers Albritton [1970] is by no means alone in holding just the opposite view, viz., that Wittgenstein left no room for toothache in toothache language-games, no room for pain in pain language-games. But this could not be further from the truth.

"it is-we should like to say-not merely the picture of the behaviour that plays a part in the language game with the words 'he is in pain ' but also the picture of the pain..not merely the paradigm of the behaviour but also that of the pain."
(P.I., 300)

This remark is qualified by the prescription that we say not that the picture of pain is involved in the language-game but that the image of pain is involved. The qualification may seem less than transparent as it stands and introduces a degree of confusion to the English version of remark 300 as a whole. There is an alternative translation of remarks 300-1 however and a lucid unfolding of their philosophical content in Peter Winch's "Wittgenstein: Picture and Representation" [1987]. In his paper Winch takes issue with

Anscombe's rendering of "vorstellung" as "image" which is rightly held to have "overpoweringly pictorial associations" (p. 76). Winch's point is that, as "bild" is usually given as "picture", if "vorstellung" is then given as "image", with overtones of "icon", then any contrast between the two terms must be lost and with that loss Wittgenstein's whole point disappears. In place of:

"It is true that in a sense the image of pain enters the language-game; but not as a picture... An image is not a picture, but a picture can correspond to it."

Winch [1987] substitutes:

"It is true that in a sense the representation of pain enters the language-game; but not in the form of a picture... A representation is not a picture, but a picture can correspond to it."

The point is then that the way pain is exemplified in a context is something to which a picture can correspond in that one can form a picture of the behaviour involved. That same picture of the behaviour will also correspond equally to acting or to mimicry or deceit even though what is exemplified, the representation, is different. Therefore, representation and picture contrast exactly as thinking of something and picturing the object of the thought contrast. The point that the representation is not itself pictorial is also well made. With an old man in a blanket Michelangelo represented God and the majesty of his love but if a picture requires some degree of genuine similarity then that representation is hardly pictorial.

Winch attempts to make sense of Wittgenstein's first claim that although the representation of pain enters the language-game it does not do so pictorially; that is, pain is represented in the context and in behaviour, linguistic and otherwise but not as a picture

(what would be a picture of pure naked pain ?) In addition Winch shows how plausible Wittgenstein's second claim is that although a picture can be taken to correspond to a representation it does not follow that the representation itself is pictorial. Earlier in his discussion Winch alludes to remarks 424-6 of the Investigations where Wittgenstein distinguishes pictures we freely choose, which suggest their own application and are readily so applicable, from those which are not chosen but seem to be presupposed by a way of talking. Talk about other minds, Winch suggests, belongs to the latter category. Such talk characteristically involves the actual inapplicability of the picture suggested and seems to leave us with only an indirect route; with second best: "while I was talking to him I did not know what was going on in his head". The mistake, Winch argues, consists either in focussing only on this picture at the expense of the particular use it springs from, or, in taking the way a person's mind is represented as itself literally pictorial. But to say this much is not yet to deny that such language does involve pictures; neither is it to eliminate or reduce them. Wittgenstein, Winch argues, is explicitly non-reductionist and emphasises that: "The picture is there" (P.I., 424).

In "Wittgenstein: Picture and Representation" Winch offers a highly plausible alternative translation with an explanatory power which dwarfs the original. Winch therefore succeeds to an extent in "Trying to Make Sense" of remarks 300-1 by demonstrating more clearly the philosophical eating here and, as regards the alternative translation proposed, that is the proof of the pudding. Despite this fact, the representation of pain remains pictorial at



least in the sense that it is ultimately a picture of pain-behaviour which is the means of representation, as it were. But if it is a picture purely of behaviour which remains then the job of refuting the crypto-behaviourist interpretation has not yet been done. However, if in describing the picture of the behaviour we cannot but use pain-vocabulary, mentalistic vocabulary, then in one sense (the non-reductionist one) Wittgenstein could not be a behaviourist. This is not made clear by Winch however. A clearer general account of this issue is given by John McDowell in his "Criteria, Defeasibility and Knowledge" [1982] and I discuss the contents of that paper below in Chapter 6.

That Wittgenstein was not a behaviourist is made clear in, among others, remark 307. Here Wittgenstein explicitly considers the behaviourist accusation, accuses himself of holding everything other than behaviour a fiction and asserts that if he is speaking of a fiction it is of a grammatical fiction. There is for Wittgenstein adequate room for pains in pain language-games, and pain is certainly represented in those games. The proviso consists in the fact that the grammar of those games be properly understood for grammar is to inform us both of the nature and the role of sensations and images. This in no way amounts to a denial of the existence of mental accompaniments or of a reduction of such to behaviour but is in typical style a plea for correct understanding. Whatever grammar is appropriate here it cannot be one of object and designation. This picture and its ensuing perplexities have been eliminated. Now however we lack a positive philosophical account of matters and the temptation is to ask what then is an image and what a sensation ? But these are precisely the wrong questions. For what

will you count as an answer beyond what has already been rejected ? The relevant terms are still employed as substantives and these interrogatives invite a description of somethings in the sense just dispensed with. What is proposed instead is a conceptual investigation via a perspicuous representation of actual word-use. Characteristically, Wittgenstein argues that the question and his proposal both ask for a word to be explained but the former leads us to anticipate the wrong kind of answer.

I have tried to show that in the hands of both the Classic Empiricist and the Cartesian Rationalist philosophical psychology has come to grief. But does the Wittgensteinean alternative presuppose a conception of philosophical psychology which reduces that discipline to a branch of morphology ? Certainly, the methodological prescription is to begin from an examination of the grammatical but the implications of that enquiry need not be confined to the linguistic; grammatical investigation can have consequences exceeding austere morphological interest. Dummett's 'metaphor thesis', viz., the view that metaphysical disputes are ultimately linguistic disputes, provides one example of a way in which the consequences of investigations into language and meaning might have ontological implications. [1963] This, I argue, constitutes a valuable rule of methodology even if the philosophical character of the disputants is not quite as uniformly dichotomized, invariant and generalizable as Dummett might suppose.

The privacy embodied in psychological concepts, the unimpeachability of the subject, is not a function of the type of

object an image or sensation is but is a reflection of the nature of the linguistic games we play with them. To maintain that only the subject can know his own sensation is at once to confuse a grammatical fact with an empirical fact. Ordinarily we frequently do know when other people are in pain. The point is that while it makes sense to doubt as regards others it makes no sense to doubt in my own case. The sceptical restriction on the applicability of "know" to the first person is itself a reflection of the grammatical fact that there is no role for doubt in that game. The privacy of sensation amounts then to saying that this is how the game is played. Grammar permits a role for doubt in the one case and rules it out in the other. That is why: "the proposition 'sensations are private' is comparable to 'one plays patience by oneself'" (P.I., 248). The philosophical prejudice of the proposition distorts the reflection of the grammatical fact that this is how the game is played.

In a sense then the paradox is resolved once the account of sensations as somethings is rejected. The point is not that sensations are therefore nothings. Clearly, Wittgenstein is not denying their existence. Nor is it that sensations are something in between the two. Rather, the point is that the something/nothing distinction, the physical-object/designation grammar is not the grammar which actually governs our use of sensation-language. To take the inappropriate grammar and impose it where it does not belong is a philosophical temptation which is responsible for many difficulties according to Wittgenstein and if anything is denied or forbidden it is this. Once that temptation is removed the paradox is removed along with it.

To argue that it is only of what behaves like a human being

that we say it has sensation is not to argue that it is only in behaviour that sensation consists. I argue from the evidence above that Wittgenstein did not identify sensation with behaviour, deny the former or reduce the former to the latter. Sensations, bound up with their representations, clearly do have a role. What is denied, and surely rightly, is a number of erroneous theories of sensation.

IV CRITERIA AND EMOTION

Analysis and clarification of the concept of emotion is an aspect of the study of human nature which is well documented in the literature of both philosophy and psychology. Traditional thought on this subject has frequently focussed upon forms of what might appropriately be termed 'feeling-theory' which identifies emotion with the presence in consciousness of particular felt qualities, analogous to sensory qualities, which are properly accessible only via introspection. Descartes [1981], for example, describes the emotions or passions of the soul as introspective perceptions, specific states of consciousness; and again in Hume the passions are described as impressions of reflection. Another form of feeling-theory has also been upheld by William James in his "Principles Of Psychology" [1950] where emotions, and indeed consciousness, are identified with bodily sensations. According to such viewpoints emotion terms are most naturally construed simply as names designating feelings which are considered explanatory of behaviour in specifying its cause as a given inner experience.

This account of emotions is not entirely an unproblematic one. It can be argued for a variety of reasons, that the language of emotion is misconceived as a set of names. The Wittgensteinian

critique of similar views concerning the set of psychological concepts points up the fact that no term functioning as a name for a state of consciousness construed as a logically private object could plausibly be said to have any shared intersubjective meaning in the context of a public language. But the language of emotion, for example, is plainly meaningful and this is just a *Reductio* of feeling-theory, as I have been arguing above. In addition it is natural to reflect that there is in fact no set of distinctly demarcated feelings to which emotion terms could correspond. Feelings of annoyance, for example, may be wholly indistinguishable from those of indignation. Similar difficulties may arise at the level of demarcation of emotional feelings from non-emotional states. Again, as is familiar from Ryle [1949] the notion of privileged access via introspection might be disputed in this context in view of the fact that a person can be wrong about his own emotions, failing, for example, to recognize a jealous state. The attribution of a naming-function to every emotion term in ordinary language and discourse would appear to ignore that sense or use of such terms in which descriptions are given of long-term dispositions to act, for example, fear of snakes or craving for oysters.

As is well-known these difficulties and others led authors such as Ryle [1949] to seek a more adequate account in a dispositional theory of the emotions. According to such views emotions are identical with dispositions or tendencies to particular behaviour patterns and explain actions not as occult causes but by reference to types of disposition expressed in law-like

propositions. Obvious problems arising from such approaches included the possibility of being sincerely afraid in the absence of any overt behavioural manifestations whatsoever. There is a price to be paid for not trying to cut the middle path, as Wittgenstein did for sensations, between somethings and nothings; regardless of which side we take.

With this much background I turn now to the role of criteria in the context of the emotions. It has been argued against Ryle-type theorizing that while behavioural dispositions may provide evidence for statements attributing specific emotions it remains possible, logically and empirically, for behaviour descriptions to be true when the attribution of emotion is false; and equally it has been argued that the converse also holds. The conclusion drawn from this state of affairs is that the two are independent in terms of truth and falsity and therefore cannot be semantically equivalent. It only follows from that fact that Ryle cannot have his identification, his assertion of logical or metaphysical identity. From the point of view of criteria no such identification was ever alleged and it is open to the criterialist to argue that the problematic truth-value discrepancy is just the condition of defeasibility which he explicitly recognizes. Of course the onus is on the criterialist to make sense of this claim and to provide it with philosophical substantiation. In fact I will defend the claim as it stands but am here concerned only to argue that if the criterial basis is a sound one then it is one from which a plausible story about criteria in the context of emotion logically follows. This context will clearly make more rigorous demands of criteria in terms of adequacy conditions. Criteria for assertions of emotion cannot be allowed to

be exclusively behavioural but must include elements of social context. Thus while certain behaviour patterns are equally consistent with, for example, both shame and embarrassment it would be wholly illegitimate to assert shame in the absence of an object for the emotion which engenders actual or perceived reprehensibility. The point is not a new one. It is Errol Bedford's [1965-57] and the importance of extending the criteria beyond behaviour narrowly interpreted can be seen by uncovering what is right, and what is wrong, in that author's account.

Theorists such as Bedford [1965-57] have identified the cause of an emotion with the object of that emotion and have analysed emotion as an evaluational response to its object. This must have some element of truth in it just in so far as evaluation is often a prominent feature of emotional states and because knowledge of an agent's evaluational response can be a good reason to attribute a certain emotion to that agent. Despite this, evaluation alone cannot provide a complete analysis for it seems that what makes emotions emotional, the felt character of the experience, might be wholly left out of account. In the entry under "Emotion" in the "Encyclopaedia of Philosophy" William Alston argues that emotions can be understood as 'bodily perturbations'. These might then be construed as the effects arising from the evaluation of the object. But again any account of why cold, rational judgement should cause such perturbations is just what has been left out. Although there is often an evaluational aspect to the felt character of an emotion the former cannot therefore constitute the latter. Therefore, there can be no plausible identification of emotion and evaluation. But

this conclusion is entirely consonant with the account of psychological terms which I have outlined as the later Wittgenstein's and which I believe to be fundamentally correct. Despite this state of affairs, the object, the evaluational response to the object and the behaviour of the subject do all express the emotion. That the object is evaluated in a certain way will characterise the nature of the emotional state. Different evaluations will represent different emotional states. In this way evaluation of the object provides a means of identifying (perhaps fallibly) an emotional state in a subject and readily functions as one criterion, among others, for emotion.

Although I am not presenting any sort of theory of the emotions here I hope my remarks show the special relevance both of the object and the subject's evaluation of it as criteria of some significance in the context of the emotions.

CHAPTER THREE

INTRODUCTION

The principal text of the Wittgenstein corpus remains the Philosophical Investigations. Of all the later texts and manuscripts under scrutiny among Wittgenstein's commentators it alone was intended for publication by its author. In examining Wittgenstein's use of "criterion" in that text I take it as already established that Wittgenstein did not use that term to give content to a reductionist form of behaviourism. This leaves two widely-argued claims as regards criteria in the Investigations. Firstly, that it constitutes the basis of Wittgenstein's own Neo-Constructivist Theory of Meaning and secondly, that the notion of a criterion is the key to Wittgenstein's refutation of scepticism about other minds. In the section below I examine the former claim. The section subsequent to that deals with the latter. Finally, in the third section I consider the use of On Certainty among commentators who defend a criterion-theoretic interpretation of Wittgenstein's later philosophy and challenge the validity of any appeal to that text for those purposes.

I THEORY OF MEANING

At this point I will only discuss the question of whether or not Wittgenstein advocated a Theory of Meaning in the Philosophical Investigations. The wider and more interesting question of the nature and potential value of the Theory of Meaning in contemporary philosophy of language I will examine in Chapter Four. In the present context however I raise the question of Wittgenstein's

attitude to theoretical approaches to the concept of meaning in order to refute a particular type of interpretation of that authors remarks.

A very clear example of the type of commentary which I wish to oppose in the present chapter is contained in Baker and Hacker's joint Critical Notice of Wittgenstein's 'Philosophical Grammar' in Mind, vol. 85 [1976]. The views expressed therein are no longer upheld by those authors but are nonetheless exemplary of a type of commentary which has enjoyed considerable currency. In their Critical Notice Baker and Hacker [1976] maintain both that:

"Wittgenstein's.. later philosophy is unified by its commitment to constructivism.."
(p.270)

and that:

"The explicit verificationism of the Bemerkungen was not a temporary aberration.. but the first step in a sustained exploration of forms of constructivist semantics."
(p.270)

The Philosophical Investigations was felt to represent the culmination of that exploration in a Neo-Constructivist Theory of Meaning erected upon a foundational Theory of Criteria. But is this plausible ? As is well known with respect to the Investigations Wittgenstein described many things as constitutive of meaning. In remark 353, for example, it is asserted that appeal to verification-procedures may enter into an account of meaning:

"asking whether and how a proposition can be verified is only one way of asking 'how d'you mean' ? the answer is a contribution to the grammar of the proposition."

In remark 560 ordinary explanations of meaning are said to be respectable accounts of meaning:

"the meaning of a word is what is explained by the explanation of the meaning ."

and, famously, on page 220e Wittgenstein advises:

"let the use of words teach you their meaning."

As regards names, in remark 43, Wittgenstein contends that pointing to a bearer can contribute to an explanation of meaning. Over pages 214-216 of part II Wittgenstein clearly maintains that expressive intonation can cause a word to be filled with meaning. This is a figurative use of the term but it is not a figure which is arbitrarily chosen, rather it forces itself upon us and:

"the figurative use of a word cannot conflict with the original one."

Rather than endorsing any single core notion Wittgenstein appears to outline a number factors which can be involved as regards meaning and its explanation. Further, In remark 65 it is contended that there is no one single feature or phenomenon which unites language. This, I argue, must be consistent with, and reflective of, the view that there is no single, unifying aspect of use which can be singled out as the central one in terms of which every other could be explained. These thoughts echo the characterisation of "meaning" in the Blue Book as an "odd-job" word to be contrasted with the idea of a word with a single, precise function in terms of which it is invariably to be understood. The same sentiment is again explicit in Remark 64 of On Certainty:

"Compare the meaning of a word with the 'function' of an official and 'different meanings' with 'different functions'."

In Remark 560 of the Investigations Wittgenstein asserts:

"if you want to understand a word look for what are called explanations of meaning."

It will be a truism to say that as a matter of linguistic fact explanations do not come in only one form or serve only one kind of purpose. Explanations will be as diverse in aim and type as the uses of the utterances they seek to explain. In so far as a Theory of Meaning strives to locate any one feature essential to semantics, be it criteria or anything else, there can, for Wittgenstein, never be a valid Theory of Meaning; for the simple reason that no such unifying factor exists. On this point, I argue, he is surely right. Even if it is contended that we call all these "meaningful" and, therefore, they must have something in common, it remains invalid to infer that they therefore all are so called in virtue of a particular feature which every single case shares. That would require independent demonstration.

It has been argued that despite the variety of factors involved in giving an account of meaning which are explicitly recognized by Wittgenstein in the Philosophical Investigations the slogan that 'meaning is use' is, at least potentially, a unifying factor capable of providing a basis for a central notion in the Theory of Meaning. Michael Dummett [ie. 1963/73], for example, has taken this suggestion very seriously. Against that idea however, I would again argue that there just is no one particular feature or aspect of use which can be singled out as the central one in terms of which all other aspects might be explained. Dummett's response here has been to streamline Wittgenstein's claims, focussing on the essentially public character of meaning. As regards the logical constants, taken as paradigm cases, Dummett picks out two aspects of use: justification for use, the 'introduction rules', and secondly, the consequences of

use, the 'elimination rules'. The condition that the reason for asserting must harmonize with the consequences of so doing produces a non-holistic account with revisionist potential. But, plausible as this sounds, it does not establish that Wittgenstein did in fact support a Neo-Constructivist Theory based upon a core notion of criteria. Even if it were the case that the 'meaning as use' slogan had great unifying power it would remain a good question why use was to be cashed out, always and only, in terms of criteria; especially given that many sentences do not have criteria. Of course, the Dummettian approach would probably take a notion of assertability-conditions or the like as the core notion here and would classify criteria as a type of such. Dummett's approach clearly merits consideration in its own right but I will consider that matter in the subsequent chapter.

The Neo-Constructivist interpretation is used here only as an example of the attribution of a THEORY of meaning to the Wittgenstein of the Investigations. If that particular interpretation could be shown to be erroneous it would still be possible to offer some other Theory of Meaning as being 'really' Wittgenstein's. But there is a more general reason why that too would be erroneous. That is, I think it can be deduced from a number of Wittgenstein's remarks that he did not adhere to any THEORY of Meaning and was in fact opposed not to some theories but to the theoretical approach. If that can be shown to be the case it must follow not only that it is a mistake to attribute a Neo-Constructivist Theory of Meaning to Wittgenstein but that it is a mistake to attribute to the Wittgenstein of the Investigations any Theory of Meaning.

I would appeal here to Wittgenstein's conception of the nature of philosophy. In his "Wittgenstein on the Nature of Philosophy" [1984] A. Kenny identifies one of the key features of Wittgenstein's conception of philosophy during the post-Tractatus period as emphasising a contrast between the descriptive and the theoretical. Description ought to be illuminating and representative. The theoretical approach is explanatory and hypothetical. The sole purpose of philosophy resides in description facilitating release from the grip of linguistic bewitchment. Language, through its grammatical categories almost suggests theoretical approaches; through noun and adjective to substance and property, for example. Not only are there metaphors but sets or structures of metaphors pertaining to concepts like time, for example, which suggest forward direction: the arrow of time, the flow of the river of time etc. In one sense it is of the nature of metaphor to mislead. As a device which explicitly identifies that which is never actually identical it involves identifying one thing with something else which, ultimately, it is not. But Philosophy, for the Wittgenstein of the Investigations, is intended to constitute a means of resisting just such temptations and suggestions by elucidating, and expanding our view of patterns of use, by description. It is, by definition, non-theoretical and non-hypothetical.

This distinction between the descriptive and the theoretical might be thought to connect with a similar concern in empirical theory. In that context the distinction between the purely descriptive and the theoretical is often denied. But both in the

present context and where empirical theories are obtainable it is a mistake, I argue, to deny the distinction. To suggest that all language is theoretical is surely to deny the possibility of mere description. Contemporary philosophers of science frequently motivate the rejection of the empiricist conception of a foundational empirical basis by appeal to the theory-laden nature of all language. But the fact that phenomenalist language, the language of experience, is parasitic upon physical-object language is what entails that the empiricist in his foundationalism has got his logical priorities wrong. It is by no means a logical consequence of that fact that when we ask for description we are always asking for, or getting, a theory.

According to Wittgenstein attempts at explanation ought to be replaced with lucid description. This is dissonant with the very possibility of Theory of Meaning as traditionally conceived. These sentiments are very clearly expressed in remark 109, again in remarks 125-6 and in remark 127: "the work of the philosopher consists in assembling reminders for a particular purpose." In these remarks Wittgenstein wholly recasts the purpose of philosophy which becomes therapeutic; a way of dissolving perplexities by actual case-studies in language whose upshot will be reports so unobjectionable as to end debate. Contra Kenny's claim this is hardly 'in the same methodological spirit as that of Descartes'.

These descriptive pictures of the actual workings of language which remain of philosophical theory will be no Theory of Meaning but will retain a point or purpose and a "light" of their own. They will enable the theorist to comprehend the various roles of "meaning". But it is surely valid to infer from these remarks, which apply quite generally to Philosophy, that it is plainly

inconsistent to describe the author of these remarks as someone who, in actual fact, subscribes to an account of meaning which identifies one, single feature, the criterion, as the definitive core notion of that theory. It remains an independent question however, to what extent any systematic, but purely descriptive, classification fixing on some, perhaps observable, aspect of use, internal to a language, or to certain language-games, could be compatible with Wittgenstein's thought. But it does at least seem clear that a grand unified theoretical story specifying content from the outside and perhaps with predictive power must be wholly incompatible with Wittgenstein's later thought. Again, questions as to the nature and value of theories of meaning must retain an importance beyond the interpretation of Wittgenstein's later philosophy but I consider those questions in detail in the next chapter.

Interestingly, it might still be possible to talk here about 'theory'; though not in the currently orthodox sense. Etymologically, the particular conception of the nature of theory I have in mind does remain faithful to the original Greek usage of the term 'theoria', the root of our own term, which relates to the verb 'to see' and to elucidating and clarifying; notions which survive, even on the Wittgensteinian conception.

It is important to qualify my position here and hopefully to clarify it. I do acknowledge that Wittgenstein used the term "criterion" in the Investigations but cannot accept that this amounts to a full-blown theory of any kind. In particular, I cannot

accept that it amounts to a Theory of Meaning. To look for anything in the Investigations remotely similar to those criterial theories surveyed by Lycan [1974] must be to search for chimeras; at best it will result in an exercise in misdescription. Wittgenstein did introduce the term into philosophical discourse and is undoubtedly responsible for the currency which the term enjoys but I do not accept that more was intended by it than its introduction in pages 24-25 of the Blue Book makes explicit. Wittgenstein's role in the matter of contemporary debate may well be one of origination and inspiration but that does not justify the claim that he held anything like a contemporary theory. It is either this inspirational role which, through esteem, ends up exaggerated to the extent of attributing to that author a "criterial semantics", or the attempt to exact a misplaced legitimation by authority. Greater weight might be given to Wittgenstein's own hope, expressed in the preface, that his work not spare others the trouble of thinking but inspire them to thoughts of their own.

Finally, it seems clear to me that Wittgenstein would have conceived of the motivation behind the desire to produce a Theory of Meaning as a prime example of the philosophers' 'craving for generality' (Bl.Bk, p.18). It is the result of such a craving, Wittgenstein argues, that we cannot accept the method of investigating philosophical problems, questions about meaning or the nature and workings of language, in terms of language-games. Further, the 'craving for generality' can be redescribed as a ^{attitude towards} 'contemptuous _^ the particular case' (Bl.B., p.17).

Interestingly, Wittgenstein includes the following remark in his discussion of the craving for generality in the Blue Book,

"I want to say here that it is never our job to reduce anything to anything, or to explain anything. Philosophy really is 'purely descriptive'." (Bl.B, p.18)

Further, I argue, the philosophical methodology which is advocated in the Investigations is clearly one of elucidation by description and not of hypothetical explanation. What is not descriptive in the Investigations is either itself a tool for description [ie. language-game] or is a plea for description as the business of philosophy. Of course, it could be argued that "criterion" is a tool for description and might be widely applicable as a crucial feature of use and a non-theoretical unifying concept. I do not think Wittgenstein would deny that, in some cases, examining criteria could help to explain how we use words as we do. Nor do I deny that we could assemble these descriptions and classify them. But this is a very long way from the type of Theory of Meaning which, for example, Baker and Hacker [1976] wanted to see in Wittgenstein's later work. In conclusion, it seems prudent to take Wittgenstein at his word in this if there is to be any possibility of appreciating the significance of his thought. I argue then, that it is not simply that in the Investigations Wittgenstein fails to propound a certain type of semantic theory but that the theoretical approach to meaning, as explanation rather than description, is entirely absent.

II OTHER MINDS

I have already outlined one approach to this problem in the Investigations when considering remark 293. There Wittgenstein

assumes a type of philosophical psychology and semantic theory which readily gives rise to scepticism about other minds. Next he points up certain bizarre consequences arising from that assumption and finally rejects the conception. This, I argue, is intended to be a Reductio Ad Absurdum of the type of position that scepticism about other minds depends upon; of its metaphysical and, in a broad sense, semantical presuppositions. There is a more positive side to Wittgenstein's account of the matter but it is not one which will necessitate a Theory of Criteria.

Scepticism about other minds is discussed in remarks 281-90, though consideration of that topic is also evident elsewhere (293, 350, for example). The nub of Wittgenstein's argument in remarks 281-90 concerns what might be described as the primacy of human being vocabulary in ordinary discourse. The picture is that the concept of a human being involves an implicit recognition of the existence of other minds as a fundamental aspect of its meaning. To correctly employ the concept at all is to acknowledge that the objects to which it applies are exactly the sort of thing which sees, hears, hurts, thinks and so on. Without this implicit recognition the concept will have been radically altered. To reject that recognition will be to reject our very linguistic training and will symbolize not a philosophical assessment but a simple objection to convention. This is eminently in keeping with Wittgensteinean methodology: not to build a philosophical bridge over yawning metaphysical chasms but to close the gap by preventing the division arising. In fact, on my reading of Wittgenstein at this point his view is very similar to the one spelled out by Strawson [1959] in his "Individuals". As such Wittgenstein's account will, of course, be open to similar

objections. It may be acknowledged, for example, that we are unable to employ mental vocabulary without being able to utilise the concept of a person. Further, that concept does not necessarily apply uniquely to me. However, it could be argued that all we can validly infer from this is that it is possible that there are other minds, not that the existence of other minds can be proved.

Much of what Wittgenstein objects to in remarks 281-90, I argue, is the formulation of the sceptical problem. The pertinent question would be: human beings as opposed to what? What sort of thing would it be that might in all respects be like a human being and yet be devoid of sensations, perceptions, even consciousness? In remark 420 Wittgenstein suggests that to seriously conceive of children as automata will result either in the words becoming meaningless or in the production of an 'uncanny' feeling. It is also implied that such a conception of other human beings could not be maintained in our ordinary social intercourse. What Wittgenstein does here is to reorientate the epistemically primary from so-called colourless bodily movements to human actions. It is a reorientation of the logical priorities of Cartesian epistemology. The strategy is to prevent the sceptic splitting bodily movements which can be known directly from human beings which can only be the objects of self-knowledge not of other-knowledge. It is not bodies that we know in knowing human beings, or at least not only bodies. It is absurd, Wittgenstein argues, to say that bodies feel pain, that it is a hand that suffers rather than a person who is aware of pain in his hand:

"one does not comfort the hand but the sufferer."
(P.I., 286).

The very ability to conceive the world in the vocabulary of

the sceptic is considered to be contingent upon the pre-mastery of a more basic vocabulary and the prior possession of certain primordial concepts, including that of a human being. The sceptic puts the cart before the horse when he invites us to envisage persons as automata or possible automata, for the ability so to conceive is parasitic upon a prior recognition of the concept of a human being and its application. The point here is a point of logical priority. The sceptical picture of witnessing bodily movements and dubiously inferring mentality becomes untenable when it is seen that if this were the case it would be inexplicable why mentality should ever be inferred in the first place. Alternately, having begun such a practice why should it be confined to living beings ? Bodily movements divorced from human beings cannot provide the foundations of epistemology. What Wittgenstein takes as basic is the very concept of a human being and its application. The implicit contents of these concepts are construed broadly:

"my attitude towards him is an attitude towards a soul
I am not of the opinion that he has a soul."
(P.I., Pt.II, p.178)

This remark reinforces the point made in 281-90. The remark itself has been identified by commentators as Aristotelian. The echo of the Greek conception of the soul is plainly audible. It is not that we infer the presence of a spectral entity, if that is not a category mistake, but that concepts like that of a human being already contain a recognition of the fact that what they are about is living beings, with all that this entails. To speak of a human being is not to speak of what is, or might be, inanimate but of what has life. To say in that sense then that a human being does not have

a soul is to say that it is dead. This is why, in Plato for instance, although the nature of the soul is up for grabs the possession of souls by human beings is not [ie. Phaedo, for example]. In "Eine Einstellung zur Seele" [1987] Peter Winch argues that the attitude, the "Einstellung", to which Wittgenstein refers is only intelligible at all in a social context. The attitude is manifested in the reactions of human beings to each other as human beings. In the Investigations, Part II, Section 4 Wittgenstein contrasts "attitude" with "opinion". The contrast intended, Winch holds, is that between being and not being reason-based. The attitude is represented in instinctive reaction which needs no justification from any rationalistic epistemology. It therefore makes no sense to ask quite generally why we pity sufferers because that is not something grounded in reason or inference, rather it is an instinctive reaction which is basic here. It follows that recognition of the social character of existence is a presupposition of the attitude of human beings one to another, a point which supports my own conclusion. In the last analysis Wittgenstein's basic argument with the sceptic would be to hold that the formulation of the sceptical position is nonsensical, or, is at least, in Hintikka's [1962] phrase 'a performative inconsistency' as it challenges what is implicit in the very terms in which it is stated. The point might be put as being that while possibly meaningful, what the sceptic says cannot be true or false at all, unless its false.

At the outset of Chapter Two I argued that Wittgenstein's attitude to the sceptic is, invariably, to construe the sceptical position as the product of misunderstanding and to argue against it

by Reductio Ad Absurdum. The attitude to scepticism exemplified in these two cases provides a thread of continuity throughout Wittgenstein's work both during and after the Tractatus period. Scepticism is described in the Tractatus as "not irrefutable but obviously nonsensical" (T-L-P, 6.51). In the Blue Book (p.58-9) the kind of extreme scepticism which leads to solipsism cannot be solved by answering it with common-sense but by "curing the temptation to attack common-sense". I have noted similar strategies in the Philosophical Investigations at 293 and 281-90 and this is further endorsed at 464 and 481. However, the theory of criteria is no part of this idiosyncratic approach in any of the cases examined.

In the case in point it might be argued that the rejection of scepticism as involving an abrogation of the meanings of the words we use is itself a criterial approach and is therefore evidence of a Theory of Criteria at work in the Investigations. But this implicitness is not described by Wittgenstein in terms of criteria. That term occurs twice during the passages mentioned [288/290]. The first asserts that if the usual language-game is "abrogated" then a criterion of identity for sensations is required. The second denies that we use criteria to identify our own sensations. That Wittgenstein did think that there were criteria in the usual case is made clear, for example, in remark 269:

"Let us remember that there are criteria in a man's behaviour for the fact that he does not understand a word: that it means nothing to him, that he can do nothing with it. And criteria for his 'thinking that he understands', attaching some meaning to the word but not the right one. And, lastly, criteria for his understanding the word right. In the second case one might speak of a subjective understanding. And sounds which no one else understands but which I 'appear to understand' might be called a private language."

Clearly, criteria are involved here, but not as a means of refuting scepticism. Criteria provide ways of telling whether or not someone understands a word. There is also, of course, the famous remark 580 which asserts that: "An 'inner process' stands in need of outward criteria". This would also exemplify the existence of criteria in the usual case and might even be thought to endorse the view that Wittgenstein held a Theory of Criteria. The single sentence constitutes the whole remark however which compounds the difficulties of exegesis and the fact that "inner process" occurs in indirect speech ought to urge caution and prudence in declaring this remark the cornerstone of any such theory. The point might be argued that Wittgenstein intended "inner process" in a sense in which he would only have discredited. As it stands the purpose of the remark is less than completely perspicuous. Remark 404, on personal identity, clearly does endorse the two points that there are criteria in the usual case and that it is not by means of criteria that I identify my own sensations or even myself:

"What does it mean to know who is in pain ? It means, for example, to know which man in this room is in pain: for instance, that it is the one who is sitting over there, or the one who is standing in that corner, the tall one with the fair hair, and so on.-What am I getting at ? At the fact that there is a great variety of criteria for personal 'identity'. Now which of them determines my saying that 'I' am in pain ? None."

Again, in remark 377 Wittgenstein asks:

"...What is the criterion for the sameness of two images ? What is the criterion for the redness of an image ? For me, when it is someone else's image: what he says and does. For myself, when it is my image: nothing. And what goes for "red" also goes for "same"."

In the normal case, therefore, there are criteria for the

identity of sensations and further I do not identify my own sensations by criteria, but that implies no refutation of scepticism.

Although the problem of other minds is not in question in these remarks it is clear that we presuppose the existence of the criteria, the criterial story is prior to scepticism and is not adumbrated to refute scepticism. If the sceptic rejects the ordinary criteria he will just lack criteria and will enter the realm of nonsense or the attempt to construct a private language, which will come to much the same.

Scepticism is the product of misunderstanding language but not just misunderstanding the criteria. To return to Winch's point: it is the *einstellungen* which is basic or primitive and that is explicitly contrasted with the theoretical or the reason-based. For these reasons scepticism is not refuted in the Investigations by criteria but is shown to be nonsensical, the product of misunderstanding language, a part of which might involve misunderstanding criteria. Therefore, although Wittgenstein did think that there were criteria in the normal case it is not by appeal to them that he refutes scepticism about other minds. Although later supporters of criterial theory may well have exploited and developed arguments featured in Wittgenstein's work this alone cannot prove that that work itself contains a Theory of Criteria used to refute scepticism about other minds, even if it may fairly be said to have inspired some.

III ON CERTAINTY

Criteria are often described as a form of evidence or grounds

for assertion which, when fulfilled, not only warrant the assertion but guarantee the truth of the assertion. Further, recognition of the satisfaction of criteria for an assertion, it is often held, can constitute knowledge of the truth of that assertion, and this despite the defeasibility of any assertion warranted by means of criteria ie. Lycan [1971], Hacker [1972], Baker [1974], Baker and Hacker [1986a]. Clearly, these are not uncontroversial claims and serious doubt has, of course, been cast on the coherence of any such concept in epistemological contexts by Wright [1984] in his "Second Thoughts About Criteria". I consider the very challenging argument of that paper in Chapter 6 but for the moment my primary interest is confined to the interpretation of Wittgenstein's later philosophy and to the allegation that a theory about criteria in the sense outlined is integral to that philosophy.

The notions of doubt and knowledge receive detailed examination, above all perhaps, in the very last of Wittgenstein's writings *On Certainty* [1950-51]. If the view which attributes such a theory to the later Wittgenstein is correct then it would not seem unreasonable to anticipate some account of the concept of criteria in that work, particularly in the light of the controversial claims made on behalf of the notion as regards knowledge. Further, such an expectation might seem the more reasonable in the light of the fact that *On Certainty* is regularly referred to by a number of commentators in their exposition of criterial themes ie. Hacker [1972], Bloor [1984]. In fact however no reference can be made to Wittgenstein's use of the term "criterion" in that work, simply because the term does

not occur once in the text of *On Certainty*. It is therefore legitimate to ask what it is that justifies reference to that text to explain Wittgenstein's use of the term "criterion" when Wittgenstein does not use that term anywhere in the text. The answer to this question is provided by Hacker [1972] in terms of the expression "grounds" which occurs frequently in the text and which is said to enjoy a special logical relation to "criterion".

The nature of the relation in question is explained by what Hacker [1972] terms "The Criterial Argument":

"The sense of a sentence is determined by its criteria. The criteria for a sentence consist of the non-inductive grounds for its application or use.."
(p.263)

The thesis is not given any textual reference at this point but is said to "lie beneath the surface of Wittgenstein's remarks". In the final chapter, "The Problem of Criteria", in a section entitled "Some Logical Features of The Criterial Relation" (p.287), a textual reference is given:

"In the Notes for Lectures he [Wittgenstein] argues that the meaning of blindness is given by specifying the criteria for blindness which are constituted by the grounds for calling someone blind."
(NFL. p.285)

In the "Notes For Lectures" Wittgenstein uses the term "criterion" once on the page referred to, as follows,

"What is our criterion for blindness ? A certain kind of behaviour."

But the term "grounds" does not feature anywhere on that page and hence the relationship between that term and the term "criterion" remains basically unexplained. Therefore, no real textual basis is given for the claim made in the Criterial Argument and from the fact that there is no use of criterion in *On*

Certainty it follows that the argument cannot be presented there. Worse still, it is clear that there are, for Wittgenstein, different kinds of ground and kinds of grounds other than criteria, for example, symptoms. The fact that Wittgenstein does not actually use the term "criterion" must cast doubt on the idea that by "ground" Wittgenstein meant "criterion". The former is clearly a more general term which might include, but cannot be exhausted by, the latter. Finally, there is another, trivial, reason why criteria are not grounds, namely, just because criteria are observable circumstances, not sentences.

What then does the Criterial Argument actually come to ? It cannot be the claim that criterial relations require grounds. I have argued earlier [above, Chapter 1] that in the Blue Book's introduction the bedrock that we hit is not a ground but a convention. Criteria may determine what is to count as a ground but that does not imply that the criterial relations are grounded. In that sense the criterial relations are groundless. Equally, it cannot be that everything which is a ground is a criterion. That would abolish the distinction between symptom and criterion which was the *raison d'être* of the terminology in the Blue Book. Importantly, the grounds mentioned in the Criterial Argument are described as "non-inductive".

Let us now consider that description in the light of the point of the Criterial argument: It allows On Certainty to be analysed in terms of criteria via the notion of a ground. But, *ex hypothesi*, not every ground is a criterion. That will be so only for uses of "non-inductive" grounds. But if that is Hacker's argument then we

are back to square one: there is no use of "criterion" in *On Certainty* but neither is there any use of "non-inductive ground". This difficulty cannot be circumvented by introducing the idea of a grammatical connection. The claim would then be that any ground which also reflects a grammatical connection will be criterial. But is that so? In "Scepticism, Rules and Language" Hacker [1986a] is clear that it is not. Criteria are a subset of the set of internal (grammatical) connections and are of a more complex order [for more detail see below, Chapter Five]. How then can the Criterial Argument be of exegetical value if it presupposes that we can recognize when a ground is criterial in advance of a coherent distinction between grounds which are and grounds which are not criterial. That distinction is not familiar either from the Philosophical Investigations or the preliminary sketches for such. The Criterial Argument, I argue, far from lying beneath the surface of Wittgenstein's remarks, compounds an error in recognising a second theoretical, hypothetical and meaning-theoretic term therein. It is a part of the frame through which the remarks are viewed, not a hidden part of the remarks themselves.

The Criterial Argument is, at least, misleading in so far as it suggests that in *On Certainty* it is criteria which confer sense on a proposition. In the actual text Wittgenstein frequently describes the language-game as a meaning-conferrer. The context of an utterance determines its meaning (228-9, 348). The language-game sets the limit to meaningful interaction. As a methodological tool for philosophy the fundamental advantage of that concept lies in its incorporation of social structure, context and in sharpening focus on aspects of linguistic practice. That concept is clearly broader

than that of a criterion. While the latter might be an important aspect of the former I would be extremely reluctant to say that it ultimately exhausted the concept of a language-game; if for no other reason than the fact that numerous language-games proceed in the absence of criteria. Not every language-game contains criteria, ways of knowing or roles for doubt or justification. Unless the rules allow for these moves then to try to so move is as ludicrous as appealing to the rules of draughts to make a move in chess.

In Hacker's picture there exists a criterially-supported realm of utterances in which recognition of the satisfaction of criteria can confer truth, falsity, certainty and knowledge. This realm is above all a grounded realm. But is this a plausible account of On Certainty ? Certainly the term 'grounds' features in that work but there can be no synonymy between that term and the term criterion as Hacker intends it to be understood. In this text, as in the Investigations, Wittgenstein is at pains to emphasise the groundlessness of our believings (166). At 307 and 387 it is emphasised that we can be certain in the absence of any grounds and that the locutions 'not certain', 'as good as certain' and 'I cant doubt it' are all perfectly meaningful even where no grounds are available. Propositions about our own experience and even the law of induction are also held to require no grounds. Therefore when Wittgenstein speaks of language-games as meaning-conferring he cannot be speaking only of criteria. If he were vast tracts of perfectly meaningful language would become meaningless. This is again spelled out in remarks 559 and 560 where it is held that the language-game is not based on grounds and is neither reasonable or

unreasonable but like life "is there". The point might be put as that while justification of particular reasonings occur due to standards and criteria of correctness within a mode of reasoning layed down in a language-game there can be no criteria of correctness, no justification of a whole mode of reasoning, external to the practice itself. In the remark that the language-game "is there", with its air of return to the definition of what is mystical in the Tractatus (T-L-P, 6.44), Wittgenstein is clearly casting doubt on the idea that the concept of grounds provides the best approach to language-games or is specially basic to them. In so far as he rejects that idea, I argue, he rejects the spirit of Hacker's whole approach. Grounds apply only in contexts, in language-games, which accomodate them. These are only a subset of the set of contexts; a fragment of all possible contexts. There is no indication that at the bottom of every language-game, as a condition of meaningfulness, there are criteria. To fail to mention that subset to which grounds are essential might be a failure of description. But to abstract that subset and impose it over the whole of language, to take it as a reductive class, is a philosophical sleight of hand which does no justice to Wittgenstein.

As an interpretative tool with which to carve up the text of On Certainty criteria leave a great deal to be desired. But it is not in terms of the contribution to a criterial interpretation of Wittgenstein that the importance of the text inheres. The work does contain, for example, much that anticipates contemporary debates and it is in this, I argue, that its importance consists. The text exemplifies a development, a broadening of descriptive horizons. From description, in the main, of particular games in the

Investigations there is movement at some points in *On Certainty* towards meta-descriptions of systems of games. Further, the twin concepts 'hinge-proposition' and 'world picture' develop an idea which is not entirely dissimilar to, indeed seems to anticipate, the Quinean [1953] notion of a 'conceptual scheme' which features in so many contemporary debates. The Quinean picture is of a core set of propositions beleived to be true and strongly depended upon. This is distinguished from a complementary set of peripheral propositions which are in some sense less than fundamental. Shared belief in a common core-set is shared admission to a conceptual scheme.

Incompatible cores represent distinct conceptual schemes,

"The totality of our so-called knowledge or beliefs, from the most casual matters of geography and history to the profoundest laws of atomic physics or even of pure mathematics and logic is a man-made fabric which impinges on experience only along the edges... A conflict with experience at the periphery occasions readjustments in the interior of the field..the.. logical laws being in turn simply certain further elements of the field."
(p.42)

The Wittgensteinean picture is perhaps less readily explicable but in some ways there are clear approximations. The core of propositions might be deemed Wittgenstein's world picture but only with some reservations. It is not that the propositions comprising the world picture are explicitly believed to be true. They may never be questioned at all (167). They precede questions of truth and falsity in order to provide a basis from which other propositions are up for grabs, in terms of truth and falsity (94, 159, 162). It is not that it would be impossible to test or examine any particular component propostion in the light of others, only that it would be unusual (163, 164). The world picture is not a hypothesis but a

linguistic inheritance acquired through training (167). It is acquired as a system and as a frame of reference. What would be impossible, because of its senselessness and lack of meaning-conferring context, would be to test large subsets of the system. When learned the world picture propositions harden into rules. These form bedrock. Empirical propositions, in terms of the analogy, are fluid and flow over this river bed. However, the bank comprises both rock and sand. Ultimately, there is not a clear division between the rock and the waters above. But there is enough of a difference to be important (94-100). Hinge-propositions form the world picture, the river bed. These are the hinges on which the debate swings, the axis around which the periphery revolves. What is a hinge in one game can be tested in the light of a hinge in another. Something must hold fast, but it need not hold universally. But this is true only of some hinge-propositions, not of all. Some are considered more basic than others. All are not equally testable.

Wittgenstein's point emphasising the lack of any clear division here does echo many of Quine's [1953] remarks, for example, that it is,

"...folly to seek a boundary between synthetic statements which hold contingently on experience, and analytic statements which hold come what may..by the same token no statement is immune to revision."
[p. 43]

or that,

"The issue over there being classes seems more a question of convenient conceptual scheme; the issue over there being centaurs, or brick houses on Elm Street, seems more a question of fact. But I have been urging that this difference is only one of degree."
[p. 46]

There is further evidence that Wittgenstein's thinking at that

time did indeed anticipate Quine's [1953] point; notably from the Remarks On Colour which is largely contemporary with On Certainty. The following passage is from Part 1, which Anscombe, the editor, dates around March 1951:

"Sentences are often used on the borderline between logic and the empirical, so that their meaning changes back and forth and they count now as expressions of norms, now as expressions of experience. (For it is certainly not an accompanying mental phenomenon- this is how we imagine thoughts- but the use, which distinguishes the logical from the empirical one.)"

Of course, I am not suggesting that the analogy is a perfect one, nor am I denying that there are important differences between these authors. Neither do I pretend to have offered a completely comprehensive account of that text; though I do hope to have highlighted some interesting issues therein. My point has only been to try to show how far removed Wittgenstein's concerns are in that text from the construction of a Theory of Meaning whose core notion is that of a criterion.

In On Certainty Wittgenstein makes clear that in order to play the same language-game speakers must know and share that which stands fast in that game. To know what the game means is to know the preconditions of play. The propositions which constitute those preconditions will be held fast in that game, yet in a different game the truth of the same proposition may well not be settled in advance of the procedures for establishing truth and falsity which the game lays down. I do not contend that the positions outlined in this text are uncontroversial, clearly they do require development and defence. I do argue though that it is more in virtue of these positions that the text is of value and not in its, remarkably poor,

potential interpretability as a defence of a Theory of Meaning with a core notion of criteria. Again, no doubt, there are certain features of Wittgenstein's philosophy as presented in *On Certainty* which have been endorsed by later theorists of criteria but that of itself does not show that Wittgenstein himself held any such view and if such theorists do wish to maintain such a view the fact must be faced that one of the prime pieces of textual evidence for 'Wittgenstein's Theory of Criteria' is a work which does not even mention that term.

CHAPTER FOUR

INTRODUCTION.

In the preceding chapters I have addressed the issue of the origins of contemporary theories of criteria. Those origins are universally acknowledged by commentators, supporters and critics alike to have their ultimate source in the later philosophy of Ludwig Wittgenstein ie. Baker [1974], Hacker [1972], Albritton [1970], Lycan [1971] and McDowell [1982] to cite but a sample. In so far as I have been wholly concerned to date with establishing the validity of the claim that Wittgenstein did not outline any such theory in those later works and is properly understood to have had only an inspirational role in this matter [above Chapter Two] so far have I been able to avoid explicit reference to anything but the most general features of contemporary theory in very schematic sketches. The fact remains that the conclusion to this debate will itself say nothing about the nature or coherence of that theory. Whether a full-blown theory of criteria can justifiably be attributed to Wittgenstein or whether, if my arguments be accepted, no such attribution is justified is of little consequence to the theory itself. The fact of this interpretational matter will neither prove nor refute the correctness of the theory; nor will it settle the purpose or role it might have in contemporary Philosophy of Language. For that reason it is now incumbent upon me to offer answers to just these questions. The role and place of the theory in the present philosophical scheme of things takes precedence over its precise contours in the order of exposition.

I THE PHILOSOPHY OF LANGUAGE

It is often argued, for instance by Anthony Flew [1979, p. viii] that an eminently sensible approach to the inquisitive newcomers' typical questions: what is philosophy ? what is metaphysics ? what is moral philosophy ? what is the philosophy of....? is to send the inquirer to the seminal texts of the particular field which is the subject of the inquiry. Teachers of philosophy will surely be familiar with this strategy from discourse with intending students. As a method it may be advantageous in that it can often do more justice to genuine questions than brief, hastily formulated definitions which may well be unduly limited by time, partisan or just fail to identify what it is in the spirit of the subject which has captivated the inquirers' interest. This methodology would at least force the inquirer to examine an actual text before dismissing (or continuing) the matter rather than simply halting that inquiry if the description given is found unpalatable, irrelevant to the motive of the question (be it sound or unsound), or just uninteresting. Suppose then that someone had asked: what is the philosophy of language ? and further suppose that in the interests of contemporariness a modern-minded teacher directed his questioner to any of the works of Dummett or Davidson or even to the Oxford Readings volume on language [Searle. ed. 1974]. In these works the questioner would find that to a large extent the entire field revolved around one cardinal hinge, namely the question: how are we to explain the fact that from exposure to a finite stock of more or less grammatically well-formed utterances the language-learner quickly comes to understand a potentially infinite number of utterances ? Accounting for this phenomenon has become a sine qua non not only of the Philosophy of Language but also of much Linguistics. In fact, concern with this question cuts across a number of academic boundaries and is debated by many of the most prominent and

influential figures in each field. Dummett [1978], for example, gives great philosophical weight to just this question and sees one particular fact as being central to the whole of the Philosophy of Language, namely, the fact that:

"..anyone who has a mastery of any given language is able to understand an infinity of sentences of that language, an infinity which is, of course, principally composed of sentences which he has never heard before.." (p.451)

Similarly, from the "Philosophy of Language", Chomsky [Searle ed., 1974] asserts:

"The most striking aspect of linguistic competence is what we may call the creativity of language, that is, the speaker's ability to produce new sentences, sentences that are immediately understood by other speakers although they bear no physical resemblance to sentences which are familiar ..the fundamental importance of this creative aspect of language has been recognized since the seventeenth century.." (p.74)

In his contribution to the "Symposium on Innate Ideas" this attitude is again exemplified [Searle ed., 1974]:

"Compared with the number of sentences a child can produce or interpret with ease the number of seconds in a lifetime is ridiculously small" (p.122)

The assertion that a weighty predominance of the edifice of contemporary Philosophy of Language rests heavily upon this fundamental concern is surely an uncontroversial one. A priori, there is no reason why this should not be seen as a highly pertinent question to which the Philosophy of Language must provide an answer. The aforementioned intending student, and indeed experienced philosophers, might therefore be forgiven for proceeding to reason in the following way: The existence of the phenomenon is undeniable. The solution to it is neither obvious nor trivial. What kind of explanation is appropriate to that phenomenon ?

Again there is no reason a priori why reasoning of this kind is unsound nor the question which is its upshot misconceived. Yet, on further examination, the student will be misled if he expects to find that debate of his reflectively-framed question dominates the thought of these authors. In fact, as Baker and Hacker [1986] point out, debate of this question: "what kind of explanation is appropriate to the fact that from exposure to the finite sample a child comes to understand a potentially infinite number of unfamiliar utterances ?" constitutes only a tiny fragment of the combined literatures of Linguistics and the Philosophy of Language while factional rivalry within one ideological conception of explanation comprises the rest. The form which the answer to that question must take is already a presupposition of the debate, not a part of its agenda. Again I turn to the works of Dummett and of Chomsky which clearly bear this out. Dummett [1978], as mentioned more briefly above, asserts that:

"The fact that anyone who has a mastery of any given language is able to understand an infinity of sentences of that language, an infinity which is of course principally composed of sentences which he has never heard before, is one emphasised not only by the modern school of linguists headed by Chomsky but also by Wittgenstein himself and this fact can hardly be explained otherwise than by supposing that each speaker has an implicit grasp of a number of general principles governing the use in sentences of words of the language."
(p.451)

Chomsky's attitude is equally presumptive:

"A distinction must be made between what the speaker of a language knows implicitly (....competence) and what he does (his performance)..... Performance provides evidence for the investigation of competence..a primary interest in competence entails no disregard for the facts of performance. On the contrary it is difficult to see how performance can be seriously studied except on the basis of an explicit theory of the competence that underlies it."
(p.73)

This attitude is equally prevalent in the views of some contemporary

philosophers of mind. It is integral, for example, to Paul Churchland's account of "Matter and Consciousness" [1984] which points up nicely the extent to which such a view is presupposed not only in thinking about language but in thinking about thinking. Above all for Churchland the fundamental concern of the Philosophy of Mind is the study of "conscious intelligence". My complaint lies not with Churchland's demarcation of the field in those terms but in the fact that, again, the explanation of conscious intelligence given is not always one that is argued for by Churchland but is more often one that is presupposed. Consider, for example, a few of Churchland's introductory remarks on Artificial Intelligence:

"These rules [of arithmetic] you already know..So you already possess a self conscious command of one formal system. And given that you can think at all, you also have at least some tacit command of the general logic of propositions as well, which is another formal system. What is more interesting is that any formal system can be automated."
(p. 100)

Why does Churchland say that unconscious knowledge of a formal, mathematical system is a precondition of conscious thought ? No argument is given to substantiate this claim, though it is far from self evidently true. It is, of course, quite clear why Churchland holds such a view given that author's attitude to Chomsky:

"..these artificial languages [BASIC, PASCAL etc] are much simpler in structure and content than human natural language, but the differences may be differences only of degree,..the theoretical work of Noam Chomsky and the generative grammar approach to linguistics have done a great deal to explain the human capacity for language-use in terms that invite simulation by computer."
(p 16)

There is a hidden premise in Churchland's first point about Artificial Intelligence, namely, that thinking is always thinking linguistically. Again, Churchland doesn't feel he needs to provide

argument for this conclusion. It is just presupposed. But once that is conceded Churchland's reasoning obviously runs as follows: if to think is to think in a language and we only learn a language in virtue of something like Chomsky's rules which we know innately then unless we had such innate knowledge language, and consequently thought, would be impossible. Although this is already a considerable presupposition I will make clear in the course of this chapter that Churchland in fact presupposes more and that it is in virtue of these presuppositions rather than any actual argument that he reaches the conclusion he wants about Artificial Intelligence.

With specific reference to the work of Chomsky however, H. Putnam [Searle ed., 1974] characterizes one Chomskian 'mode of argument' well:

"..there is the 'argument' that runs 'what else could account for language learning ?' The task is so incredibly complex that it would be miraculous if one tenth of the human race accomplished it without innate assistance. (This is like Marx's proof of the Labour Theory of Value ..which runs, what else could account for the fact that commodities have different value except the fact that the labour-content is different)
(p.133)

As indicated above this 'mode of argument' is present not only in the writings of Chomsky but also in those of Dummett. Putnam [Searle ed., 1974] is rightly unhappy about what is an extraordinary form of argument. Not least because it is no form of argument at all but is a form of rhetoric. It is not even a genuine question but a rhetorical one in the mouth of Chomsky. My inability to conceive of alternative forms of explanation other than my own has no implications whatsoever for the possibility of there being valid alternative conception : this must be because the earth is flat, what else could explain it ? and so on ad nauseam. Putnam does not bring fully to light what is unhappy about this whole matter and himself has no developed alternative although he is not

prepared to nail the lid tight shut on traditional learning theory. Baker and Hacker [1986] do present an alternative (furthermore one which naturally accomodates a central role for a criterial thesis although these authors do not make anything at all of this potential) which I shall discuss later in the present chapter.

II THE SCIENCE OF LINGUISTICS: A HYPOTHETICO-DEDUCTIVE MODEL.

Even at this early point in the discussion it already becomes difficult to maintain an account of matters which is sufficiently general to keep the cited authors under one and the same umbrella of enterprise without running the risk of blurring the differences of detail (they are, I argue, differences of detail, not of principle). The question I want to put at this juncture is as follows: what kind of explanation is being presupposed by Dummett and by Chomsky ? Because the answer to this question may not be identical in each case and in order to avoid the charge of misrepresentation I shall examine this question with reference to each author in turn.

Chomsky frequently reiterates that the formal theory of grammar which is ultimately going to provide the answer to the question raised by the 'creativity of language' is itself an empirical hypothesis. [Searle ed., 1974 p.122, 125, for example]. There may however be a divergence between Chomsky and Dummett on the correct answer to the question of whether the empirical hypothesis is, or is not, an hypothesis which asserts the existence of certain hidden mechanisms whose operations, characterized by the formal theory, constitute that understanding in virtue of which speakers actually master their language. However, I would argue that only an affirmative answer to that question could actually

provide a solution to the creativity problem as conceived by either author. It is clear at least that Chomsky is so committed to the concrete existence of such mechanisms in the strong sense (of a mental 'ontology') outlined. This is made perfectly obvious by Chomsky [Searle ed., 1974] in a number of passages and in a footnote (p 72. 4.) where Chomsky corrects A. Riechling's [1961] interpretation of his work:

"[Riechling] asserts that obviously I could not 'be said to sympathise with such a "mentalistic monster" as the "innere Sprachform" '. But in fact the work that he is discussing is quite explicitly and self-consciously mentalistic (...that is it is an attempt to construct a theory of mental processes)."

"...a grammar is an account of competence. It..attempts to account for the ability of a speaker to understand an arbitrary sentence of his language..if it is a linguistic grammar it aims to discover and exhibit the mechanisms that make this achievement possible."
(Searle ed., 1974, p 73)

Chomsky's Theory of Transformational Generative Grammar is, therefore, an empirical hypothesis which both asserts the existence of innate mental mechanisms which make understanding a language possible and attempts to describe the operations of those mechanisms, to describe those mental processes in which human understanding consists. That description, if correct, (it is held always to be empirically evaluable) will literally make explicit just how in fact the phenomenon of the creativity of language is possible. Again I reiterate that my concern, like that of the intending student, lies not in painting in the mechanical fine detail but in ascertaining just what kind of explanation is being given. At first sight Chomsky's mode of reasoning (if empirical at all) is surely empirical in a strong scientific sense. It is most readily subsumed under the hypothetico-deductive model familiar from the

writings of Karl Popper [1963, for example], just in so far as it conjectures imaginatively the existence of entities and their functions in terms of which certain phenomena are to be explained and leaves itself open to empirical refutation by the evaluational criterion of 'fit' with actual language-use. Given this situation is it now fair to say that the model of explanation employed by Chomsky is the theoretical physicist's hypothetico-deductive model which postulates entities and describes their functions ? Plainly not. For Popper [1963, esp. Ch.1] the whole *raison d'être* of the methodology of conjecture and refutation consists in falsification. What provides systematic constraints, guarantees objectivity and ultimately, therefore, constitutes rationality on that model is explicit testability. In a strong sense made clear by Popper [1963] falsification makes scientific knowledge possible and therefore makes scientific progress possible. By virtue of its role as demarcation-criterion with respect to the scientific what is not testable and genuinely falsifiable just is not scientific. Of course, it will be argued that Chomsky [Searle ed., 1974] has provided empirical evaluation criteria by means of which refutations may be effected. But that criterion is only of application to any given grammar of a given realization of the ultimately innate mechanism. The innateness hypothesis is not refuted when a grammar or grammars is refuted; only the individual grammar is. In Kuhn's [1964] terms the scientist not the paradigm is being tested. The innateness hypothesis is not itself a testable hypothesis. The evaluation criteria do not apply to it, for the very existence of performance is evidence for competence. Every time a child learns a language the innateness hypothesis is confirmed. There is no test, crucial, conclusive or otherwise for the concrete existence of the structure-independant *faculté de langage*. John Lyons [1971], who is far

from an ardent critic of Chomsky, acknowledges just the lack of scientific testability involved:

"Chomsky ..is committed to the view that "if an artificial language were constructed which violated some of these general principles, then it would not be learned at all.. (The Listener, 30th May, 1968 p.688). But this hypothesis, as Chomsky's critics have pointed out, is not subject to direct empirical verification...Nor is it at all clear how one would go about designing an acceptable psychological experiment bearing less directly on the issues involved." (p.111)

In a later footnote Lyons [1971] cites a response by Chomsky to this charge:

"He..[Chomsky]..suggests that I should point out that .. it is generally accepted by modern empiricists 'that meaningful hypotheses, in general, must only meet the condition that some possible evidence have some bearing on them" (p.114)

It should be pointed out here however that not all empiricists are logical positivists, as Chomsky seems to presuppose. It might also be noted that not all of Chomsky's critics are empiricists in any sense and that the only 'possible evidence' seems to be the phenomenon of the creativity of language which, when seen in Chomskian terms, can but confirm Chomsky's view. The point is not that the innateness hypothesis is meaningless because it lacks testability. It is rather that in so far as these hypotheses lack testability so far are they unscientific. In short, there is nothing which could count as a refutation of that hypothesis, empirical or otherwise. But in these terms the Chomskian Theory of Understanding, far from being a genuinely scientific hypothetico-deductive conjecture, is essentially a Conventionalist stratagem a la Popper [1963 esp.Ch1].

I would also suggest that if we return to the particular grammar at hand, namely Chomsky's, it would be found that no area of performance

could even be seen, conclusively, not to fit. Not least because there is no fully articulated Theory which any piece of evidence could be seen not to fit in such a way as to produce a refutation rather than an artful extension or re-drawing of boundaries. But the philosophy of artful extension is the philosophy of the conventionalist twist; not of conjecture and refutation.

It can only be concluded therefore that the type of explanation presented by Chomsky is not as straightforwardly scientific as it might first appear. This fact prevents, for the moment, any further question being raised as to whether or not the type of explanation being offered is adequate to the explanation of the original phenomenon; that is a point to which I shall return.

However, I would be prepared to argue that the innateness hypothesis functions as the Lakatosian [1978] 'hard core' of the Chomskian research programme, or, in Kuhn's [1964] terms, defines the Chomskian paradigm. If we accept that scientific theories are more articulated, composed of more distinct parts, than Popper [1963] could comfortably allow, then the innateness hypothesis might be seen to be the frame for explanation, that which makes explanation possible. As such it would be held fast such that it was not itself up for grabs in terms of truth-value but determined what was so appraisable. The innateness hypothesis would then be a very different mode of description from the empirical, metaphysics rather than science. In these terms it might still be allowed that the accumulation of anomalies might bring about change at this level in the structure (it is clear that many contemporary linguists might want to see Chomsky as a necessary step on the way, even if the way is not clearly signposted)

III THE MACHINE IN THE GHOST

I will now consider a characterization of the type of explanation Chomsky is proposing which is not unrelated to the type just considered above and which, I believe, lies at the heart of Chomsky's thesis. Before outlining my interpretation it will be useful to consider a few more of Chomsky's remarks:

"..it makes sense to say.. that you and I know English, for example, that this knowledge is in part shared among us and represented somehow in our minds, ultimately in our brains, in structures that we can hope to characterize abstractly, and in principle quite correctly, in terms of physical mechanisms"
[Chomsky 1980, p.5]

Again I refer to Chomsky's [1971] reply to Riechling cited earlier [above, section 2] and to Chomsky's [1980, p.185] analogy between the language faculty and a "physical organ". It seems quite clear that what Chomsky is offering as explanation to the creativity of language is in fact a causal explanation. When the hearer hears a novel utterance this is processed by the alleged mechanisms referred to which are conjointly sufficient to produce understanding. More properly, on exposure to initial finite stock innate, structure-independent principles will conjointly produce a grammar for the given language. Next, on hearing a novel utterance, that utterance will be processed by the rules and principles of the formed grammar in various ways. The effect of this processing will be precisely an understanding of the meaning of that utterance. The repeated iteration of that process will make actual the possibility of understanding an infinite number of novel utterances; at least in principle ie. in the case of the Ideal Speaker-Hearer. In his "Semantic Theory and Tacit Knowledge" [1981] Gareth Evans is clear on this point, as clear, at least, as the very notion of tacit semantic

knowledge allows him to be. Evans' paper is itself a reply to Crispin Wright's "Rule-Following, Objectivity And The Theory of Meaning" [1981]. Wright considers three views of tacit knowledge; the third of which interests us here. This 'third option' defends a strong sense of tacit semantic knowledge adequate to representing the compositionality constraint on semantic theory for, it is held,

"..all the interest of the theories or sub-theories which have been constructed lies in their capacity to exhibit the meanings of complex expressions as a function of the meanings of their parts."
(p.118)

This feature is described as "structure-reflection" by Prof. Wright. Wright [1981] is critical and, rightly, sceptical about the idea of a 'strong sense of tacit knowledge' which the third view defends and Evans [1981] notes that,

"It would be unfair to complain that Prof. Wright did not make this third option terribly clear, for it is one of his points that it is not very clear."
(p.120)

Despite this obvious constraint Evans goes on to make lucid one possible account of a tacitly known semantic theory capable of representing the structure-reflecting characteristic in terms of dispositions to judge utterances as tokens of sentence-types with relevant, associated truth- conditions. Of the notion of disposition employed in his account Evans remarks:

".it is essential that the notion of a disposition used in these formulations be understood in a full-blooded sense. These statements of tacit knowledge must not be regarded as simple statements of regularity, for if they were anyone who correctly judged the meanings of complete sentences would have a tacit knowledge..When we ascribe to something the disposition to V in circumstances C, we are claiming that there is a state S which, when taken together with circumstances C, provides a causal explanation of all the episodes of the subject's V-ing (in C)....The decisive way to decide which model is correct is by providing a causal, presumably neurophysiologically based, explanation of

comprehension."
(p.s 125-7)

In all of these claims I think Evans is both perfectly articulate and perfectly correct. If the theory is really worth its salt this is exactly what it must deliver. Having said that, I think there is more to the point that the very notion resists precise formulation and I have graver doubts about this possibility, for even if such an explanation can be given let us not presuppose that it is the kind of explanation that we require, that it is even appropriate. I will make clear my doubts in the process of argument.

But first I wish to forestall an objection here: if, as is indeed the case, Chomsky is a dualist, how can the mind (which contains the language faculty) link into the causal chain? This is, of course, nothing other than the traditional mind/body problem and in the quoted remarks the lie is given to this point. Ultimately Chomsky, and Churchland, are simply identifying mind and brain, holding both equally susceptible to neurophysiological explanation. I can now make clear the extent of my complaint against Churchland. In presupposing a Chomskian account of language-acquisition Churchland has presupposed an analysis of thinking which already involves a materialist analysis of mind. It then becomes very easy for him to reach the comfortable conclusion that the programme of so-called strong Artificial Intelligence is entirely well-conceived and bound to culminate in success. All that has to be done to create a thinking, metal mind is to reproduce the Chomskian formal theory which exists in the wetware of the meat machine and automate that system in some man-made machine. My point is not yet that strong Artificial Intelligence is misconceived, but that Churchland has simply presupposed the entire story which he requires to get his view off the ground without

ever making clear to the reader that he has done so. On the interpretation of Chomsky, Lyons [1971] makes plain Chomsky's version of dualism as follows:

".. he [Chomsky] differs from Descartes and most philosophers who would normally be called 'mentalist' in that he does not subscribe to the ultimate irreducibility of the distinction between 'body' and 'mind'....he [Chomsky] makes the point that 'the whole issue of whether there's a physical basis for mental structures is a rather empty issue' because, in the development of modern science, 'the concept "physical" has been extended step by step to cover anything we understand', so that 'when we ultimately begin to understand the properties of mind, we shall simply ..extend the notion "physical" to cover these properties as well. He does not even deny that it is possible in principle to account for mental phenomena in terms of 'the physiological processes and physical processes that we now understand'."

(p.107-8)

It is precisely this ultimately reductionist version of dualism and the explanatory monism which makes the mental (eventually) explicable in physical terms which in turn makes dualism compatible with the type of explanation which, I suggest, Chomsky is offering. By parity of reasoning the appeal to dualism is rendered ineffective.

It is worth emphasising that this locates a mechanical procedure at a very deep level of consciousness. The (ultimately) neurophysiological hardwiring which constitutes the language-acquisition device and equally constitutes understanding is not available to the subject. It is not known but cognized. It is cognized in an essentially non-cognitive sense. It is UNconscious intelligence. That is, it is not in the stream of consciousness whose daily turbulences and perturbations compose a mental life. Rather it is situated at a deeper level in the geist which cannot be reached by conscious introspection. It is, to paraphrase, the machine in the ghost. I do not wish to labour with criticism of the notion of innateness. The faculty clearly fails to satisfy adequate criteria for knowledge. Simply calling this knowledge something else without amending

its role in the Theory in any way is a sleight of hand which one does not have to be a philosopher to feel suspicious about. I do not believe that Chomsky has provided any reason to reconsider Locke's classic refutation of the Cartesian brand of innateness and in either form the hypothesis is as dubious now as it was then. Again my concern lies in what kind of explanation is being given and, I argue, the role of competence is that of a causally deterministic mechanism (or group of mechanisms) with respect to performance. What makes us rightly recalcitrant to accept a causal explanation of language-use in the way I have suggested Chomsky invites us to is surely the following thought: isn't that type of explanation explicitly contrasted with the type of explanation we would want to give of correct language-use ? It is certainly in stark contrast with the type of explanation which any competent speaker would give. But the explanations of the speaker under Chomsky's regime will actually be taken only to be inductive evidence for the existence of the machine whose mechanics our ordinary explanations grope feebly towards.

The fact that language-use (performance) could be explained causally does not imply that any such explanation could satisfactorily answer the question raised by the creativity of language. Certainly, there exist neurophysiological prerequisites to language use but it is by no means as certain that reference to these facts is sufficient to explain the question^{of} meaning. To provide a purely causal story in answer to that philosophical question is surely just to have missed its point. As Waissman [1965], in another context, makes clear:

"We shall say that the meaning of a sign 'a' is the effect which it should have, not the effect that it will have. This, of course, is nothing but a contribution to the grammar of the word 'meaning' to prevent it being confounded with that of the word 'effect'."

(p.102)

I do not expect that this will convince those involved in the type of research to which I have referred. To people working within the confines of this particular ideology it will clearly require a Kuhnian gestalt-switch, a paradigm-shift, to come to see that this is so. To date I have been concerned only to raise some doubts as to the scientific status of the explanation involved and to appeal to intuitive good sense about language-use to see the inappropriateness of either type of explanation. In the section following the next I will adumbrate some philosophical argument to ground these doubts but first I shall extend my treatment from Linguistics to the Philosophy of Language.

In terms of the metaphor in the preceding section Baker and Hacker [1986] stands to the dogma of the machine in the ghost as Ryle's "The Concept of Mind" [1949] stood to the dogma of the ghost in the machine. The former work has not enjoyed the reception of the latter and is therefore considerably less influential. Despite this I accept the soundness of the critique presented there and am in broad agreement both with many of its critical arguments and with its alternative solution to the problem of creativity. I hasten to add that the arguments I deploy can be attributed to those authors only where explicitly stated; those authors cannot, of course, be held responsible for errors in my own thought as represented here and below. References to works are as quoted and are to Baker and Hacker [1986] only where stated; I do not confine myself to the sources to which those authors refer. At the end of this Chapter I present a critique of the positive views of that text (and their alleged pedigree). Finally, I emphasise that the text contains not

a single mention of a theory of criteria. It uses, but does not mention or appraise, the latter term in the ordinary way. Hence, while I include myself in the luddite spirit, as it were, Baker and Hacker cannot be blamed for any of my positive views.

IV WHAT IS A THEORY OF MEANING ?

As Baker and Hacker [1986] take pains to argue (rightly) Theorists of Meaning in contemporary Philosophy of Language are notably less enthusiastic about committing themselves to strong psychological theses than their counterparts in Linguistics. In particular the question of the psychological or physiological reality of Theories of Meaning is one for which philosophers such as Dummett and Davidson are less keen to make strong claims on behalf of their theories. What I earlier referred to as 'mental ontology' will be the focus of the current section. Here, I shall argue, Theorists of Meaning can be represented on a sliding scale of commitment, as it were, with Chomskian linguists at the top and, I argue, theorists such as Donald Davidson at the bottom. The crux of the question in which I am interested has been well put by Putnam [Searle ed 1974]:

"..the I.[innateness] H.[hypothesis] is supposed to justify the claim that what the linguist provides is 'a hypothesis about the innate intellectual equipment that a child brings to bear in language learning'. Of course, even if a language is wholly learned it is still true that linguistics 'characterizes the linguistic abilities of the native speaker', and that a grammar 'could properly be called an explanatory model of the linguistic intuition of the native speaker'. However, one could with equal truth say that a driver's manual 'characterizes the car-driving abilities of the mature driver' and that a calculus text provides an explanatory model of the calculus-intuitions of the mathematician.' Clearly it is the idea that these abilities and these intuitions are close to the human essence, so to speak, that gives linguistics its 'sex appeal', for Chomsky

at least."
(p. 132)

This is well put and applies equally to Theorists of Meaning. That is, let us suppose that the Theorist of Meaning is putting forward his formal theory (actually his formal theories ie. of sense, of force, of type-sentences as abstract objects etc.) merely as a way in which he can represent by a recursive theory how to grind out an infinite number of sentences in that formal theory each of which is assigned "truth-conditions". That is, where the notion of a truth-condition is to be understood as a technical term of that theory, say, in the Tarskian sense, in which the names of well-formed formulas of the object-language on the left hand side of the biconditional are taken into statements in the meta-language on the right hand side ie. $(\varphi) \equiv \varphi$. If this is the case then I would argue that just in so far as it is understood as a conjunction of formal devices making no claim either to apply to anything other than the formal language[s] with which it is concerned or to exist in the mental ontology of speakers so far as it is both unobjectionable and a possibly useful object for (contrastive) comparison with natural language and no doubt many practical uses in the broad field of computer science. As Putnam [Searle. ed., 1974] makes pungently clear however this is not what Chomsky is claiming. Nor, I argue, is this what Dummett is contending. If this were all that a Theory of Meaning purported to be then it would be opaque why it was thought to bear upon the problem of the creativity of language. That is, it would not be the case that the formal device constituted a Theory of Speaker's Understanding in terms of which that phenomenon could be explained. It would also be less than perspicuous that it was a theory about meaning in natural language rather than of the term 'meaning' understood as a technical term of the formal

art. In fact Dummett, at least, is making the strong and not the weak claim here. A few of that author's remarks will make this obvious:

"We..derive our knowledge of the sense of any given sentence from our previous knowledge of the senses of the words that compose it ..any theory of meaning which is unable to incorporate this point will be impotent to account for the obvious and essential fact that we can understand new sentences."
(Dummett, 1981. p.4)

"..philosophy has as its first..task, the analysis of meanings..the deeper such analysis goes, the more it is dependant upon a correct general account of meaning, a model for what the understanding of an expression consists in.."
(Dummett, 1981. p.699)

"A theory of meaning will..represent the practical ability possessed by a speaker as consisting in his grasp of a set of propositions..the speaker derives his understanding of a sentence from the meanings of its component words.....the knowledge of these propositions that is attributed to the speaker..[is] ..implicit knowledge."
(Dummett, 1976. p.70)

"..until we have, for..one..central notion for the theory of meaning, a convincing outline of the manner in which every feature of the use of a sentence can be given in terms of its meaning as specified by a recursive stipulation of the application to it of that central notion, we remain unprovided with a firm foundation for a claim to know what meaning essentially is."
(Dummett, 1976. p.137)

It is made quite explicit in the cited remarks that a formal theory of the type outlined above is implicitly known by speakers. The question remains however whether the formal theory merely represents the implicit knowledge of the speaker or whether the speaker actually possesses and 'knows' that formal theory. If the latter, psychologistic view is the case then it is from that theory, its dictionary (lexicon), axioms etc that a speaker derives his understanding. On this view, possession of the formal theory would make understanding possible for it would be the interaction of heard novel utterances with the formal mechanisms which

would produce an understanding of that utterance and, for such thinkers, unless such a theory is postulated we can neither answer the question posed by the problem of the creativity of language nor can we ever lay legitimate claim to know what meaning even is. If the theory of meaning is so construed then it seems to be a theory of understanding in the strong Chomskian sense. Even though the formal character might deviate from that paradigm to some extent, still the presence of the theory plus the hearing of the sentence looks conjointly sufficient to produce understanding of the latter. But, if this misrepresents Dummett, if the former, non-psychologistic view is Dummett's view, that is, if, according to Dummett:

"A theory of meaning of this kind is not intended as a psychological hypothesis...is not concerned to describe any inner psychological mechanisms which may account for his having those abilities. If the task of the theory of meaning is merely to give a proper analysis of what the speaker can do then why is implicit knowledge of the theory ..attributed to the speaker."
(Baker/Hacker [1986], p.340)

Further, these authors go on to ask why Dummett describes that implicit knowledge as knowledge which "... issue[s] in the ability to speak" (ibid. p.359) Dummett is caught between Scylla and Charybdis. Either he is committed to the strong theses which the cited remarks seem to imply, but to which he seems reluctant to commit himself, or he is committed to a thesis which is insufficiently strong to provide a solution to the problem for which it was designed. Unless the knowledge of the formal theory is attributed to the speaker in such a way that it can generate his understanding of the novel sentences then that theory cannot actually explain how it is that we do so understand. A similar price might also prove to be the cost of asserting such a thesis however.

V A CRITICAL APPRAISAL

The decisive battle in the war of the entire issue I have discussed in the present chapter is clearly to be fought out in the terrain of the philosophy of meaning, understanding and explanation. What runs through the philosophies of Chomsky and Dummett as a common core is the belief that understanding a language will ultimately prove to be a matter of operating a calculus which is implicitly known, in Dummett's view, or whose basic structure is, in the case of Chomsky, innately given. These derivations in the case of any given speaker will be so internalized as to be inscrutable not just to other language user's but to that speaker himself. Understanding will consist in the conjunction of these mental processes and the philosophy of understanding will consist in unearthing these processes and displaying them as a formal calculus or recursive theory. Given that what a speaker understands when he understands a sentence is the meaning of that sentence the meaning of that sentence will be identical with the product of those mental processes, the last line of the derivation as it were. Meaning, on the Chomskian model at least, will be a matter of mental process. Admittedly, that may not be Dummett's view of meaning. Perhaps meaning is to be given sentence by sentence by assertability-conditions and for mathematical statements in terms of proof procedures but nonetheless, unless the Theory of Meaning is implicitly known by speakers at the end of the day then there is no 'full-blooded' Theory of Meaning.

In Chapter Two [above] I outlined a number of arguments from Wittgenstein's Philosophical Investigations which sought to establish at least four points. Firstly, that fundamentally different concepts apply to understanding as against mental processes; regardless of how precisely

these latter are constituted. Secondly, that no mental process could provide a sufficient condition of understanding. Thirdly, that no mental process could provide a necessary condition of understanding. Finally, that it is extremely problematic how words could ever come to be used as they are if understanding a word is construed as being ultimately a matter of making mental connections; as, say, associating symbol and meaning at the atomistic level of the linguist's lexicon. It is fair to say that the upshot of this plethora of arguments is that: "...nothing is more wrong-headed than calling meaning a mental activity.." (P.I. 693). These arguments still seem cogent to me and Chomsky has said nothing to persuade me of their invalidity. It is not, perhaps, incumbent upon Dummett to argue against this body of work but it is surely not unreasonable to expect Chomsky to do so. That is, it seems clear that Dummett's view, as expressed in Dummett [1978], is that:

"Among those ideas of Wittgenstein which..relate to the main outlines that a successful theory of meaning must assume..we can..select some about which it would be universally agreed that any attempt to construct a theory of meaning must come to terms with.." (p.452)

Unfortunately, that attitude does not seem to have been shared by Chomsky. However, acceptance of the validity of Wittgenstein's arguments must be wholly incompatible with any thing but the rejection of that type of view which I have been considering as regards Chomsky. Hence, I would seem to be back at square one in terms of answering the question of creativity.

Before turning to consideration of a positive account of that question I will offer a critical appraisal of Dummett's conception and present some qualms about that question itself.

Firstly, if I may refer again to Dummett [1976]. When a paper like

this is read for the first time close attention must be paid to the terms meaning and understanding which are often used in non-standard ways ie. the question of "...what meaning essentially is" [ibid. final paragraph]. I would suggest that what is actually going on is more a matter of redefinition than clarification. Argument is given to persuade the reader to see meaning in a particular way or ways ie. as equivalent to truth-conditions or to conditions of assertability. Thus persuaded we come to accept new ways of talking about meaning, to change our use of the term, so as to see better what was really meant by it (prior to the change). In this way new uses come to be along with new criteria for explaining, justifying and criticising the use of it. This must be in part why first exposure to such work involves a certain philosophical disorientation. Taking truth and meaning as examples, we might come to see meaning as being given by truth-conditions. The question I wish to raise is whether this is a matter of discovery or of definition. Is it not that in changing our criteria for the use, justification and criticism of meaning we come to say that meaning is truth-conditions rather than to see that it is. The terms are, I argue, redefined theoretically and the newly legitimated uses have much more the character of technical terms in a philosophical theory. This is more a misgiving than an argument but in view of the entrenchment of belief in this area it seems legitimate to utilize full means of persuasion.

One of Baker and Hacker's [1986] arguments: the argument against action at a distance (in linguistic rules not in the laws of Physics) which might be felt more philosophically respectable (it is in fact a Reductio Ad Absurdum) runs as follows. Both Chomsky and Dummett speak of certain linguistic rules and principles as partly constituting a speaker's implicit knowledge. Admittedley, for Dummett, there can be no

guarantee, a priori, that the theory actually reflects the actual workings, as it were, of a speaker. Nonetheless, it seems clear that such rules never explicitly feature in speaker's explanations of meaning, nor do they feature in teaching, criticising or commending use. Further, if such rules were incorporated in any such way there is no guarantee that they could even be recognized for what they actually are. There is a sense in which these rules are never actually followed. It will be replied that they are followed implicitly. But is the notion of a linguistic rule which does not, and could not, feature in speaker's explanations etc itself a coherent notion ? Is it not the case that for a symbol to be a symbol it must be actually and publically used as a symbol and be recognisable explicitly as such ? It follows from this type of view that what is given by speakers as explanation, justification and criticism of use is not the meaning of the term used. But if what is explained in explaining the meaning of a word is not the word's meaning then what is it that is being explained ? Is it not the case that anyone who denied that what is explained when a competent speaker correctly explains the meaning of an utterance is the meaning of that utterance would usually be taken either to have misunderstood the concept of meaning or to be rejecting that conception and offering one of his own ? The upshot of these concerns is what Baker and Hacker [1986] term "the argument from logical space". The point is that there is not enough logical space within one and the same concept of meaning for both our ordinary practices with meaning which recognize that correct explanation can give meaning and the Theorist of Meaning's practice with meaning which recognizes no such thing.

I now depart from Baker and Hacker [1986]. It seems clear that the

theorist of meaning will argue that what is explained when the meaning of an utterance is explained by a competent speaker is itself simply empirical evidence for his having understood the meaning of that word; even if the constraint that explanation and use must harmonise is imposed. Alternately, it could be argued that an ability to give such explanations forms part of linguistic competence. This is perfectly correct. But if that is so why do we need to posit implicit knowledge to speakers at all ? If the question is what kind of evidence there can be for attributing knowledge of meaning to a speaker then the answer is what better evidence could there be than the ability to produce explanations, to correctly use, to criticise and to teach use ? But all this is not simply empirical evidence for someone's knowing the meaning. Rather his knowing the meaning consists in these abilities. Again the question arises why must such knowledge be implicit or tacit ? If I were to explain my saying: "He's in pain !" by invoking criteria for pain ie. I say "look he's screaming etc. He must be in pain because that's just what it is to be in pain!" is what I have said simply empirical evidence for my knowing the meaning of the term pain ? There is considerable contrast here with the modes of explanation I have been considering. In this case something public explains my understanding of the terms I use, it justifies my saying and explains the meaning of what I have said. It invokes a standard, or norm, and, as Baker and Hacker [1986 A] put it, formulates a rule. But all of this is explicit and there seems no good reason why my formulating an explicit, consciously-known rule for the use of the word "pain" through my explanation, based on the meaning-specifying criteria, should only add up to empirical evidence for my implicitly knowing other rules which I cannot produce but in which my understanding "really" consists. This is the argument from logical space.

Once ordinary, in this case criterial, explanations of meaning are admitted as explanations of meaning there is no space left for appeal to tacitly known rules. Again, I do not accept that our ability as regards compositionality need be implicit when we can actually and consciously explain why we use words in sentences as we have been taught and as we would teach others. Why must we deny that this is sufficient and postulate implicit knowledge where explicit knowledge already exists ? I would argue for conscious composition which we have been consciously taught. The rules we use must, in principle at least, be capable of being consciously formulated. Conscious composition becoming habitual, subconscious and reflexive with time and repeated use seems a far more plausible candidate than rules which are not and cannot be inferentially involved in an agents' beliefs, desires or ordinary knowings, be those rules implicit or innate.

In my view it is as a form of explanation of meaning that the value of criteria consists. That is, in a form of explanation which cites only the public, the explicit in order to teach, justify and criticise. This form of explanation can explain meaning and can indeed provide evidence for understanding. Understanding (and misunderstanding) can be obviously exemplified in this way. But does this 'evidence' simply give empirical support for understanding ? Of course, in the case I have outlined it is quite possible that I am wrong, unlike the effect of a mechanism, it may be that I have been duped by an actor. In other words, even if there are circumstances which are criterial for P and even if, in this case, those criteria in fact obtain it remains quite possible that P does not. Therefore, this type of criterial inference cannot be entailment. If P were entailed by the criteria it just would not be possible that I be

wrong about P given the criteria. But, as I have argued, that is clearly not the case.

Despite the defeasibility involved in the assertion it is clearly not possible that I was wrong about the meaning of the term I used. The criteria determine what counts as evidence for pain. The criteria for pain are part of what that term means. A term's criteria are, by definition, necessarily good evidence. In this sense criterial explanations involve explicit appeal to and thus formulate simple (informal) rules to which competent speakers are agreed. Thus the criterial relation is a meaning-relation. So the relation between my criterially-based use, my criterial explanations and my understanding the meaning of what was explained cannot be simply empirical. My argument is that knowing the meaning of "pain" just is knowing the criteria and explaining, using, criticising and teaching on that basis. Further, if that is so, what else do we need to explain my understanding of that term as a speaker.

There is a further argument in Baker and Hacker [1986] which I will adumbrate in support of the line of reasoning I have been developing. I follow the title of that text [1986] in naming the particular argument "the argument from sense and nonsense". The argument is itself of Wittgensteinian vintage and exploits the notions of context-dependence and purpose-relatedness which I have already outlined [Chapters 1 & 2 above]. The point is made that explanations of meaning are purpose-relative and that meaningfulness of utterances generally are context-dependent in the following sense. Baker and Hacker [1986] raise the question of whether a competent speaker can establish what an utterance means solely on the basis of a knowledge (implicit or explicit) of semantic rules ? The existence of the science of pragmatics in the

tripartite division of linguistics acknowledges the legitimacy of concern with context among those whom Baker and Hacker oppose. Baker and Hacker point to the hierarchical nature of the tripartite division in which pragmatics features lowest and are themselves arguing a much stronger case than that science can allow. Their thesis is that whether an utterance is sense or nonsense is determined by the context of that utterance. Terms are explained relative to paradigm cases of their proper application without which there could not be explanation (learning and teaching). The paradigms themselves are just sets of circumstances and are themselves circumstance-dependant. Therefore, whether an utterance is sense or nonsense, whether an explanation is an explanation of meaning depends equally upon the context and the purpose of the utterance or the explanation.

Independently of those authors a thought-experiment comes to mind which may support their conclusion. Imagine a grammar ruled out as nonsense a sentence which speakers consented to as meaningful outlining contexts in which it seemed so to them. Presumably this would be a ground to reject that grammar. Now imagine an utterance such that in no possible context it had sense and an utterance which could be guaranteed to make sense in every possible context. My question is: what would a Theory of Meaning rule out and conversly what would it admit in ? How could that Theory map every possible meaningful utterance in advance of use which can authorize new ways of speaking ? Again independently of those authors I make the point that the paradigms to which they refer feature circumstances which are criterially related to the terms they explain. This makes clear the role of criteria as standards, points up the fundamentally normative role they perform and emphasises their

importantly non-inductive nature.

In both Theories of Meaning and of criteria there are operative conceptions of rules, understanding and explanation. The content of these concepts vary to an extent which I think endorses the argument from logical space. It is equally plain which is the more intuitive account. Of itself that quality cannot guarantee correctness yet, I argue, if we are to transgress the bounds of the intuitive we have a right to expect a rationale to do so. I do not accept that an appropriate or sufficient reason has yet been outlined by those Theorists to whom I have referred.

The Theory of Criteria in my sense is not a theory of meaning in the sense to which Wittgenstein was opposed and neither, I will argue, is it a sufficiently generally applicable notion to constitute the core notion in a theory of meaning in the sense in which Dummett has proposed. The criterial circumstances of P are determinants of P's meaning which fix limits on its applicability. They need not express the full significance of P in every possible context, every game in which it has a role. Many utterances have no criteria but are not for that reason meaningless. Nonetheless, where there are criteria for P, these specify what P means and are not simply empirically or symptomatically related to P. In other words I do not accept that all grammars, the procedural rules or presuppositions are the same in every universe of discourse. Grammars are various and form a continuum which will not easily facilitate the identification of a Dummettian notion of such centrality as to explain meaning uniformly throughout. I argue that within the disparate forms of language very different notions might be central to understanding meaning. In the light of the diverse nature of language how could meaning be a concept which allows a precise mathematical function from a single central notion to specified uses for the entirety of a language ? The

purpose of examining particular grammars is to expose a logic which we already use. Why should it be the case that the patterns of rational thought are uniformly dichotomised in every circumstance ? Certain grammars exhibit a different pattern of reasoning and there is no reason why others might not contain still other inferential structures. It is always possible, in practice not in principle, that the inference be defeated. This is part of the context of the inference. What is not possible is that the entire pattern should be out of synchronisation in every case. The existence and practical success of the pattern refutes that claim. Justifications of particular inferences occur due to standards, criteria of correctness within a mode of reasoning. But there can be no criteria of correctness for the criteria of correctness. The mode of reasoning contains the possibility of justification but there can be no justification of the whole mode of reasoning external to the practice itself. It may not be meaningless to assert the possibility of such a breakdown but it is difficult fully to so conceive the world. If the world did so alter would we know what to say at all ? Would language even be possible without stable connections ?

In my remarks above I do not intend the term 'grammar' in the Wittgensteinean sense, if that is a completely clear sense. The sense in which the Theory of Criteria is a logic or a grammar is that in which it provides the procedural rules of the activity. It is the logic of the social institution or structure. Classical logic would be a grammar in this sense and any logic text-book would be a Theory of Grammar.

If asked what they mean speakers might give criteria in the sense I have outlined. They would not supply a recursive theory but explain in ordinary mundane ways. The evident difficulty for linguists in

describing the full empirical reality of their theory and for philosophers in establishing in a detailed way what kind of theory is appropriate is testimony to the lack of availability of such a Theory. In what sense do language user's know something which they do not and cannot bring to consciousness ? In this sense what could I not be said to know ?

VI AN ALTERNATIVE SOLUTION

Even if my critical appraisal is accepted the question of the creativity of language remains unanswered. But is this as mysterious a phenomenon as we have been led to believe ?

Firstly, I follow Baker and Hacker [1986] in challenging the idea that a child can actually understand an infinity, potential or otherwise, of novel utterances. This seems a gross exaggeration. Certainly it will understand many words and sentences but there will remain many which it will not understand, words which have to be learned, sentences which are not fully perspicuous. Do we ever stop being in this situation and possess some sort of complete linguistic understanding ? I think this is a misleading way in which to approach the matter and, like Baker and Hacker [1986], suspect the point here is more properly that the concept of number is not a helpful one in this context. This intuition can, I think, be developed by examining the value of the notion of an Ideal Speaker-Hearer. Just how helpful is that notion ? The problem is, I argue, that no Actual Speaker-Hearer is ever in a sufficiently approximate position. That is, it is surely just wrong to think that any actual speaker really does understand an infinity of novel utterances. Certainly, an actual speaker can gain an understanding of novel utterances, but this will be more of a long, hard, learned slog than is

presupposed by the type of view I have been considering. This is not to deny structure-reflection, rather it is to deny that we can make sense of the Ideal Speaker-Hearer in any useful way. I would argue that we cannot justifiably attribute to any Actual Speaker-Hearer the same sort of recursive structure, whose deductive closure will involve potential infinities, that we attribute to the Ideal Speaker-Hearer. Of course, it will be argued against this that we could bootstrap up to the Ideal from the Actual by repeatedly adding one, as it were, and thus that a Sorites Paradox arises. But despite the Sorites Paradox the theory about the understanding of the Ideal Speaker-Hearer will still not be a description of the Actual Speaker-Hearer's understanding, to which, I argue, unlike the Ideal Speaker-Hearer's understanding, a limit must apply. Further, it remains possible that an answer to the Sorites Paradox will emerge; perhaps in terms of a distinction between certain kinds of transitivity in certain contexts.

Further, it is worth noting that, even for novel utterances, the criteria of understanding remain the same. Again, I have already argued that meaning and understanding are of the nature of abilities, capacities to do things. If as Baker and Hacker [1986] suggest we pause to examine other human abilities in the way in which we examine these abilities the matter will seem ridiculous. The fact is that when we are able to ride a bike we are thereby able to recognise and ride an infinite number of bicycles but no one argues that this phenomenon requires a recursive theory as an innate mechanism. Admittedly this example, unlike a language, is non-representational, but the same point clearly holds good for painting in just that representational sense. It seems clearer here, perhaps, how inappropriate it is to apply the concept of number to an

ability; for it is precisely this extensibility in which the ability consists. In addition, the relation between an ability and its exercise is plainly not causal. It is both possible to possess an ability to do something but to fail for whatever reason to do so and to do something by luck or fluke rather than ability. Obviously, both the bicycle and painting analogies might well be objected to as lacking any analogous structure and as not being rational, at least, not in the same sense, as the ability to speak a language. In a recent review of "Language, Sense and Nonsense" in the Times Literary Supplement [Jan 11th, 1985] Crispin Wright notes that often "...focus is lost in the jumble of rhetoric and repetition..." and certainly the bicycle analogy has a rhetorical air. However, Wright's comment must be put in context with his further remark that, "Baker and Hacker have a good hand here". That is to say, despite its rhetorical air, the bicycle analogy does point to a crucial factor in the debate which requires a great deal more consideration, namely, the question of the nature of abilities. Baker and Hacker are not alone in holding that question to be vital to progress in this field. An excellent account of the importance of just such questions is given in Dummett's [1987] recent "Reply to Prawitz":

"I have for some time been dissatisfied with what I wrote earlier about understanding and knowledge...it...is essential that speech be treated as, normally, the conscious activity of a rational agent : his reasons for speaking as he does are assessed on the basis of what he is presumed to know or suppose, including what he knows or supposes to be the meanings of his words...Whether or not understanding is properly termed knowledge, it is like knowledge in providing part of the basis on which the intentions underlying utterances rest. It is probably too simple minded to say that it consists of an implicit knowledge of a theory of meaning for a language. The dichotomy theoretical/practical knowledge is too crude to be applied to knowledge of a language. If I cannot ride a bicycle I may still know what it is to ride one; but if I do not know Tibetan, I do not know what it is to speak it. Our need is therefore not to

dispute over whether knowledge of a language is genuinely knowledge, still less over whether it is theoretical or practical knowledge, but to refine our excessively coarse conceptions of theoretical and practical abilities." (p.283)

In these remarks an important emphasis is given both to the notion of understanding as an ability, and to the idea of speech as a "conscious" activity. There is no rhetorical air to these remarks which, I argue, are fundamentally correct and indicate clearly and lucidly the way forward towards a solution of what remains a deep problem. Both in the case of Baker and Hacker [1986] and Dummett [1987] there is evidence of a far better approach to the problem I have been considering and, furthermore, one which does not involve a notion of 'innate' or 'implicit' knowledge. Despite this, the (mainly Wittgensteinean) arguments which I have highlighted must represent constraints on the type of account which can ultimately be given. Those constraints may not, of course, completely preclude the construction of Theories of Meaning, though they might well cast very serious doubt on the viability of a "full-blooded" Dummettian Theory of Meaning. Again my own arguments are not the only ones which tell against that enterprise. Wright [1987], for example, offers some very challenging arguments against the "global" requirement ie. that the Theory should be neutral as to the target language. It may well be that different stories about meaning turn out to be appropriate to different areas of discourse, possibly, relative to the purpose of the enquiry and therefore it may be that Dummett's view is too essentialist, or, at least, over generalised. Further, in Chapter Six I hope to cast doubt on the view that an account of content can be given from outside the area of discourse, as it were, at least in the case of psychological language.

Finally, explanation of a concept is normative and has a social role

of citing standards of correctness. This is particularly clear where criteria are concerned. We point out "cues for use", as David Bloor would have it [1984], "ways of telling" in McGinn's terms [1984] and means of identification. But these are no more than the ordinary criteria used in teaching, learning, criticising and justifying. They determine what counts as evidence for what, when it is appropriate to say, what it means to be an X, to be in X, to have X. These ordinary rule-formulations are explicit knowns generally agreed to and obviously used. What is very different in this from Chomskian rules is precisely that if a rule explains an action the rule itself must be known to the speaker and be capable of figuring in his intentions, explanations etc. This hardly seems an unreasonable or unduly stringent demand. I conclude with Baker and Hacker that the answer to the question which the creativity of language poses: how can a speaker come to understand numerous sentences following exposure to a finite stock? is either trivial: he has mastered a language or is the request to adumbrate every skill, every ability which being a language-master involves. The latter might indeed be a promising research programme but that will not be facilitated by misconceiving from the start the job at hand. In the last analysis contemporary philosophy of language has confused the object of a legitimate and deep sense of wonder with a deeply mysterious object.

VII A COMPARISON: LANGUAGE AND STRUCTURE IN MODERN CONTINENTAL PHILOSOPHY

It has recently been argued, by Henry Staten [1985] for example, that the profoundest affinity between linguistic philosophers of the Anglo-American and Modern Continental traditions exists between

Wittgenstein and Derrida. Concern with language and structure is, of course, common to both these authors and to both traditions but I will argue that the affinity between the two authors and that between the two traditions is fundamentally misdescribed by Staten, and others, and in so doing will compare the type of view which I have discussed in the present chapter with the approach familiar from Modern Continental Philosophy. I proceed critically however from a consideration of the views of Henry Staten [1985] and Newton Garver [1973].

On page fifteen of his "Wittgenstein and Derrida" Staten explains Derrida's concept of "differance" (with an 'a') as, "temporally extended syntax". Newton Garver expresses a similar view in his preface to Derrida's "Speech and Phenomena" when he describes "differance" as:

"the phonemic structure of the sound system which provides meaning, intelligibility"
(p. xxv).

Notably, Garver also describes "grammar" as the ability to:

"...put together words in acceptable phrases and modify those words as required. The details of this requirement are studied by grammarians; and though superficial from the point of view of philosophical problems it is nonetheless a prerequisite to other sorts of competence."
(p. ix).

Both authors are concerned to establish some form of common ground between Derrida and contemporary philosophy of language and the definitions of differance quoted serve to orient our understanding of that term in a linguistic direction. Ultimately, for Staten this is a first step towards bringing the later Wittgenstein into the picture. Both therefore attempt to make that step from Derrida to the Philosophy of Language, at least partly, via a linguistic reading of "differance" and both, I argue, do so only in a confused and confusing way.

In a sense, of course, Staten and Garver are right to describe

concern with syntactic structure as common both to the Anglo-American tradition of philosophy of language and to Derrida. But it is utterly mistaken to think that there is any real point of intersection here with the later thought of Wittgenstein. Further, both Staten and Garver describe differEence and not differAance in Derrida and are wrong to think that the latter is to be found anywhere in modern Anglo-American linguistic philosophy. But Garver is most obviously mistaken about the value of grammar to philosophy.

As I have argued more than once; to a great extent the entire field of contemporary philosophy of language, and much linguistics, revolves around one cardinal hinge: answering the question posed by the so-called problem of the creativity of language. The answer to this essentially Chomskian question: how, from finite linguistic input, is infinite output (in principle) possible ? is always and invariably given in terms of just the ability which Garver describes as grammar. In addition, the answer is conceived as literally formal, a recursive theory. As many philosophers of language and numerous linguists have argued: structure-reflection, compositionality, is vital to the axioms of any theory of meaning. It is also widely held that there is some right answer to this structural question in terms of which we will cash out another right answer: the answer to Chomsky's question, once and for all, even if the precise character of either answer presently eludes us. Here then is a clear a concern with structure. But explaining language-use and language-acquisition by, and as, a formal calculus is nothing other than anathema to the later thought of Wittgenstein as, of course, is everything that Chomsky's question presupposes.

As regards structure, grammar in Garver's sense, and syntax, the point of contact between the two traditions lies not, on the Anglo-

American side, in the later philosophy of Wittgenstein but in more orthodox linguists and linguistic philosophers. Further, while many such orthodox thinkers at times pay a certain lip service to Wittgenstein the question of the creativity of language, which includes the concern with structure, is not one of those times. Furthermore, on the modern continental side, the real point of intersection, the linking of linguistic structural concerns is really not in Derrida, but in one aspect of Saussure as expounded by Derrida in his essay on "Difference" [1973].

The connection between Saussure and the contemporary philosophy of language is motivated by Saussure's conception of the proper object of study of linguistics as 'la langue' which can be taken to be directly comparable to Chomsky's notion of competence as the proper object of linguistic investigation. On this, Jonathan Culler [1976], the author of the introduction to Saussure's "Cours" and the Fontana Modern Masters volume, points out that:

"La langue, Saussure argued, must be the linguists primary concern. What he is trying to do in analysing a language is not to describe speech-acts but to determine the units and rules of combination which make up the linguistic system."
(p.30)

But how close is 'la langue' to Chomskian competence and is this in effect what contemporary philosophers of language are up to? I would argue that it is not. In taking sentences to be the products of individual choice and therefore as belonging not to la langue but to parole Saussure precludes from la langue the very stuff of which Chomskian linguistics is made. Saussurian syntax just does not include concern with regular patterns, with the rules of combination which make possible the Chomskian creativity story. Of Saussure, Chomsky himself has

said:

"He [Saussure] was quite unable to come to grips with the recursive processes underlying sentence-formation and he appears to regard sentence-formation as a part of parole rather than la langue of free and voluntary action rather than systematic rule. There is no place in his scheme for 'rule-governed creativity' of the kind involved in the ordinary everyday use of language."
[Culler, 1976 p.83]

While we can accept a broad similarity of orientation then we must also accept a profound disanalogy between, on the one hand, la langue and parole, and on the other, competence and performance. A disanalogy furthermore on the very issue which Chomsky's philosophical import depends upon.

Both contemporary philosophers of language and Saussure might well agree that, in a sense, the correct composition of differences, conceptual and phonemic, constitutes meaningfulness, though it is doubtful they would agree on much else and indeed doubtful that they would agree in the same sense. The overt Holism of Saussure stands opposed, for example, to the molecular, or atomistic, positions which are prevalent among theorists of meaning and again, while for Saussure, in some sense, it is the differences which are ultimately important, more important than the signs, it is doubtful that, say, Crispin Wright, could really, in one of his own phrases, get any sort of grip at all on that idea. As Culler [1976] asserts in a telling phrase, "...when one actually analyses a language it becomes extremely difficult to avoid speaking as if there were positive terms" (p.86)

Conversely, it seems to me that what is of interest to Derrida about language is precisely the play of differences, in the Saussurian sense, involved among the signs of the system but not, unlike Wright etc, the play involved in a particular language (English, French etc). Rather,

Derrida is interested in the play of, "traces" which makes differences possible and is therefore the condition of possibility of concepts. Because difference is the possibility of systematic signification it might be spoken of in relation to any sign system whatsoever. Indeed, Derrida discusses it not only in relation to Saussure but also in relation to Nietzsche, Levinas, Hegel and Heidegger such that all are cases of the writing most characteristic of our epoch. But this expository strategy remains, quote, a "blind tactic". Derrida denies that the play of differences can be mapped rationally in a right order for it is the 'origin' of all our terminologies and classifications and is not derived from them. Difference "opens up the space for philosophy". It is a 'point' beyond every particular structure and its reflection and as the possibility of structure it cannot be formally characterised at all. In fact, it ultimately can neither be named nor said. While the Anglo-American tradition might be said to be interested in the differences which structure particular sign-systems Derrida, it seems to me, despite likely denials, is involved in a more overtly metaphysical inquiry into the play of differences which is the possibility of structure. So there is a profound disanalogy between both traditions at the very (Chomskian) juncture where an analogy might really seem pertinent to the Anglo-American tradition and equally, what Derrida finds most appealing about Saussure is the very point at which most philosophers of language cease to want to follow. But I have conceded a broad similarity of orientation between Chomskian and Saussurian linguistics, as indeed Chomsky admits, could that similarity not provide a basis for bringing in Wittgenstein? Certainly, yes, but only as an example of someone profoundly opposed to any such investigation in principle, ie. only as someone who could not be

further removed in his attitude to language.

I will not however conclude completely negatively on this subject and I am prepared to argue for a more genuine point of contact explicitly between the writings of Derrida in "Speech and Phenomena" and those of Wittgenstein in his Philosophical Investigations, that is, a direct link unmediated by appeal to some common interest in syntax or the philosophy of language, contemporary or mediaeval. I begin from a notable and interesting stylistic point of contact between the writings of Wittgenstein and Derrida, which Garver and Staten do not note, and a sort of fundamental agreement of "einstellungen" between these two authors. This can best be seen when Derrida [1973] asks: "What is difference?" [p.145]. For even to answer that question in its own terms, is already to make a mistake, to lose philosophical depth:

"If we answered these questions..we would fall below the level we have now reached. For if we accepted the form of the question, its own sense and syntax we would have to admit that difference is derived and ordered from the starting point of a being-present".
(p. 145)

Now compare Wittgenstein's remarks that:

"a sensation..is not a something but not a nothing either"
(P.I., 304)

and further:

"One ought NOT to ask WHAT images are..but how the word 'imagination' is used. The first question also asks for a word to be explained; but makes us expect a wrong kind of answer."
(P.I., 370)

Again, to answer the question in its own terms is already to make a mistake. It is to presuppose the metaphysics of substance, or, in Derrida's case, the metaphysics of presence.

To make clear that a more important point of agreement underlies the stylistic similarity I will examine what Derrida actually says about

difference. The term is introduced as a multivalent one with at least two senses. The first is simply that of non-identity or difference. The second is of deference where what is denied is put off till later. Staten and Garver have got hold only of the first sense. Brief discussion of that sense has already involved appeal to the notion of a trace. The notion of a trace is a denial of the notion of the presence of the present and self-presence which Husserl's transcendental phenomenology depended upon. For Husserl language is a matter of indication and expression. In its indication a sign stands for something else, anticipates something not present. Indication is essential to communication but inessential to meaning Expression is the semantic content of the sign which is present only in self-present consciousness. As I understand him, Derrida's first complaint is with the idea that the imagination as the means of expression is adequate to the notion of entirely self-present meaning. In fact Derrida holds the imagination to be itself produced from modified past experience and therefore not exclusively or self-sufficiently present. This is a good point and, I argue, is equally true of conscience. The second argument challenges the very notion of presence, of a pure present moment. Any moment or trace intrinsically involves past and present as aspects of its very nature:

"the trace retains the mark of a past element and is
already marked by relation to a future element"
(p. 142)

Derrida uses the [Husserlian] terms "retention" and "protention" to mark these features. Therefore, the trace constitutes the present but that already involves past and future. What the present is is itself determined by what is not present. So the present is not simple or primordial. The present is derivative from the play of differences. This

brings us to the second (deference) sense of difference. The trace is a dynamic, self-constituting and self-dividing interval. As such it involves both spacing: the becoming spatial of time in the subdivision of the interval, and temporalization: the becoming temporal of space. Thus what for Husserl was pure meaning, the self-present moment becomes in Derrida the synthesis of traces which is the play of difference. The ramifications of such a view are striking for neither meaning nor self are ever purely present. In one sense then Derrida is denying that meaning is ever present and asserting that meaning arises only out of the play of differences within a system of signs.

It is at this juncture that the link with Wittgenstein's later thought becomes clear. Both of the remarks I have quoted are from key passages in two of Wittgenstein's Private language Arguments, I use the unusual plural deliberately here. Now, the arguments against private language form the penultimate stage in a movement, a philosophical process which defines the later works, and which I will briefly outline, necessarily in a very sketchy way. The first step is to ask: what is the meaning of a word ? this is answered only negatively as we would expect for the question is of the same type as: what is an image ?, what is a sensation ? Progress depends not on answering the question but on transforming it, which Wittgenstein does, by absorbing it into the question as to what it is to understand the meaning of a word ? This movement is clearly illustrated both in the Blue Book and in the Philosophical Investigations. The next stage, which particularly interests us here, is contained in the Investigations, so there is a nicely chronological pattern to the movement, where Wittgenstein deploys a range of arguments against the idea that meaning is a mental process and indeed against the idea that meaning is ever present to us in

consciousness, in whole or in part. The final and most famous step is the conclusion that meaning is use and thus Wittgenstein proceeds to the question of what it is to use, to follow a rule.

Like the Wittgenstein of the Investigations, Derrida argues that the process of signification is a move away from self-presence, in so far as it does not begin with self-presence but is constructed from the convention and practice of public language. This vigorous critique of the idea that meaning is ever present certainly gives us a better point of intersection between Derrida and Wittgenstein and we might see the Husserlian position as the ultimate private language which would in fact make language impossible. Having acknowledged a genuine agreement and affinity I want to return to Staten [1985], simply to refute his interpretation of Wittgenstein and in that way show that the kind of agreement that Staten finds is illusory.

Firstly, there is the question of a mythical Wittgensteinian deconstructive impulse. In "Wittgenstein Deconstructs" Staten argues that:

"Wittgenstein wanted to loosen up crystallized patterns of philosophical thought to force real thought..for which he would have no ready made answers but would have to forge new language sequences.."
(p.64)

and that:

"In the continued movement of his language the reassuring forms of usage which he has previously struck as though bedrock become impediments.. and deconstruction renews.. this is the deconstructive impulse."
(p.65)

Firstly, I argue, there are no "new language sequences". One set of jargon is not replaced by another and the Investigations is defined by the absence of what Staten [1985] describes as the first stage of

Wittgenstein's Deconstruction. Secondly, there is no second stage, for there is no 'as though' about it. We just hit bedrock in ordinary forms of reasoning and explanation (ie. Bl.B p.24, P.I., 217). Again, on p.75, Staten [1985] argues that Wittgenstein's method is one of destabilization of these reassuring forms of usage, that what is important to Wittgenstein is syntax and that his method is 'blind and mute'. But Wittgenstein does not destabilize ordinary use, nor is his method of ubersicht of use as clarification, leading to the solution of philosophical problems, in any sense, 'blind and mute'. Further, while the contemporary philosophy of language might be described as preoccupied with syntax and semantics, Wittgenstein is defined by his lack of preoccupation with such. In short, if this is the deconstructive impulse then it is a myth.

The extent to which Staten consistently misrepresents Wittgenstein is exemplified in two particular commentorial errors. Firstly, Staten (p.90-1) refers to Rush Rees' claim that Wittgenstein described himself as a disciple of Freud'. Staten admits that the extent of the influence is hard to gauge but then quotes the Rees remark, quotes Freud and leaves the implication clear. But this completely misrepresents Rees remark and Wittgenstein's attitude by taking a single phrase out of any context. Wittgenstein does say this, according to Rees, but he also says:

"Once when Wittgenstein was recounting something Freud had said and the advice he had given someone, one of us said that this advice did not seem very wise. 'Oh certainly not', said Wittgenstein. 'but wisdom is something I never would expect from Freud'."

(L & C, p.41)

Further, Wittgenstein himself is reported as saying,

"I have been going through Freud's "Interpretation of Dreams" with H. and it has made me feel how much this whole way of thinking wants combatting"

(L & C, p.50)

Further, Staten (p.95) describes Wittgensteinian "grammar" as a "joke" and: "family resemblance is a.. pun, or a rhyme". No textual reference is given to where Wittgenstein says this, and in fact the reference that is given (31), in appended notes, is to Derrida's "Margins of Philosophy". This type of commentary is shoddy, misrepresentative and completely misleads the reader. Finally, I have to draw attention to a pattern of argument in Staten [1985] which is exemplified in his interpretation of Wittgenstein. In many cases Staten considers, and offers unusual accounts of, arguments contained either in the Philosophical Investigations or On Certainty and attempts to justify highly unorthodox interpretations by reference to isolated remarks from Zettel. But while some of the remarks which compose that text can be dated by cross-reference to manuscripts numerous others simply cannot be dated and might well have been written in the 30's. It cannot be a legitimate strategy, I argue, to settle the interpretation of issues in Philosophical Investigations and On Certainty by reference to obscure remarks whose purpose simply is not clear.

As to Derrida, I argue that he can only be understood with reference to his own tradition, not to the later Wittgenstein. In Derrida's essay on Differance the deeper point is more reminiscent of Hegel than Wittgenstein and it is Hegel whom Derrida explicitly quotes and discusses immediately prior to raising the question of what differance is:

"The infinite, in this simplicity is-as a moment opposed to the self-identical -the negative. In its moments, while the infinite presents the totality to (itself) and in itself, (it is) excluding in general, the point or limit; but in this, its own (action of) negating, it relates itself immediately to the other and negates itself."
(p.144)

The play of differance always involves a postponement of something,

as we have seen, and Derrida describes this as a "crossing out" (ie. p.134) in a number of passages. This negating movement is much more reminiscent of Hegelian dialectic than the later Wittgenstein and this is endorsed by Derrida's remarks about "...the very profound affinities which difference thus written [ie. with an 'a'] has with Hegelian speech.", (p.145) and indeed, it is held, "A definite rupture with Hegelian language would make no sense" (p.145). Finally, I note Derrida's claim that "if the word 'history' did not carry with it the theme of a final repression of difference, we could say that differences alone could be historical."

However, even here there is less in common than we, and perhaps Derrida, might like to see for the word history does carry with it the theme of a final repression. Again, on p.149 Derrida asserts that:

"we shall call difference this active discord of the different forces and of the differences between forces which Nietzsche opposes to the entire system of metaphysical grammar"

But Nietzsche's megalomaniacal, heraclitean metaphysical analysis of existence in terms of a plurality of forces, locked not in a conservative, English Utilitarian and polite, struggle for survival but driven by a heinous will to power also carries with it the theme of eternal recurrence which posits some underlying order, transcendental pattern. Hence, if John Lewellyn [1987] is right in asserting only, "A Point of ALMOST Absolute Proximity to Hegel" then the same holds good of Derrida's proximity to Nietzsche.

It is appropriate to close, as Derrida himself does, on the inherent difficulty in clarifying difference which is ultimately neither word nor concept but the condition of possibility of both. Given the role of difference as a point of origin and its status beyond language, beyond

metaphysics perhaps, it is by its very nature not just difficult to express but, ex hypothesi, inexpressible. In the attempt to express the inexpressible there might be a more genuine meeting of minds. For Derrida clearly has in common with the early Wittgenstein an inability to heed the warning of the seventh proposition of the Tractatus:

"Wovon man nicht sprechen kann, darüber mus man shweigen".

CHAPTER FIVE

INTRODUCTION

Concern with Wittgenstein's use of the term "criterion" is conspicuously absent from Baker and Hacker's text "Language, Sense and Nonsense", but that was not the only publication made by those authors that year (originally 1984). In contrast, "Scepticism Rules and Language" [1986A] does contain an account of the notion of a criterion, which is given a new, but still prominent, role in Wittgenstein's later philosophy. In the present chapter I propose to critically examine the role and nature of the concept of a criterion which Baker and Hacker [1986A] attribute to Wittgenstein and to compare that conception to Crispin Wright's first thoughts about the notion of a criterion and its possible role as regards anti-realist semantics in his "Anti-Realist Semantics: the Role of Criteria" [1982]. I do not intend to offer any detailed criticism of Wright's [1982] account here, simply because Wright himself goes on to outline certain deep problems which no criterialist can afford to ignore in his own "Second Thoughts About Criteria" [1984]. My critique of Baker and Hacker's [1986A] account will culminate in highlighting a question which those authors are aware of, but do not themselves answer, and which provides the very basis of Wright's [1984] 'second thoughts', which argue that this particular question cannot be properly answered and therefore that the notion of a criterion must be rejected as incoherent. That question, together with Wright's [1984] critique generally, will be fully and critically examined in Chapter Six.

I SCEPTICISM RULES AND LANGUAGE

In what follows I cannot pretend to have offered anything more than the briefest sketch of the nature and importance of the rule-following considerations. In the present context I am concerned only to unearth and examine the concept of a criterion which Baker and Hacker's [1986A] contains. To that end I move rapidly through the early chapters and limit my account to the background necessary to understand the later chapters which contain the account with which I am primarily concerned.

Baker and Hacker's [1986A] text does not contain any explicit retraction of earlier positions, despite the fact that a great deal clearly has changed. As might be expected, given the argument of "Language, Sense and Nonsense" [1986], the term "constructivism" is not to be found in either the table of contents or the index and no attempt whatsoever is made in this work to attribute a Neo-Constructivist semantics, or any other kind of semantics, to the later Wittgenstein. Instead, Wittgenstein is taken to have identified meaning with ordinary explanations of meaning and, it is held:

"To cut meaning down to the size of accepted explanations is to eliminate the logical space in which theories of meaning move and have their meaning"
(p.118, par.1)

This claim is itself a version of an argument which I have called "the argument from logical space" [above, Ch.4] and is familiar from "Language, Sense and Nonsense". Criteria, therefore, cannot any longer be taken to provide "A New Foundation For Semantics" [Baker, 1974] in Wittgenstein's later work and this fact bears witness to a fundamental change in Baker and Hacker's attitude to the interpretation of Wittgenstein's later work.

As its title suggests "Rules, Scepticism and Language" treats mainly of rule-scepticism and is in the first instance a book about the

philosophy of following a rule. The motivation for that concern can be seen as identical with that underlying the publication of "Language, Sense and Nonsense" [1986]. In each case the authors see themselves as attempting to correct a fundamental and highly influential misunderstanding of Wittgenstein's later work. Further, while that consideration certainly provides an impetus to the publication of the work, it may also be the case there is at least one less explicit motivational factor at work. The equation of meaning with "accepted explanations" is the premise from which the argument from logical space is inferred. That premise is not an unproblematic one but is plainly essential to Baker and Hacker's [1986] position. If that position is to be a tenable one its premise must be defended. But, it is against the viability of just that premise which the force of rule-scepticism is directed.

In relevant recent literature (such as Wright [1980], Kripke [1982], for example) formulations of rule-scepticism have frequently been arrived at in the following way:

To understand what a word means is to know how to use it in appropriate contexts; to know what to do next with it. Knowing what to do next with a word is a matter of being able to follow the relevant rules of correct employment. Action and explanation testify to that ability in the individual case. At any given point of use however, a rule of potentially infinite applicability will have been applied only a finite number of times. At that point an indefinite number of interpretations of the rule will be equi-viable on the basis of past uses (the adder might be a Kripkean quadder etc.) As Wright [1980] has pointed out, there is no basis, a priori, for non-zero probability assignments among the available

interpretations. How then do we know which rule we follow, or even that our usage is rule-governed at all ?

Of course, the foregoing story involves a certain picture; a picture of the meaning of a word as a cumulative history of its past uses. Each new occasion of use is added independantly to the set of old uses. The challenge is to provide a criterion of identity for the rule governing the word. When are we following the same rule ? In fact, the argument is a radically self-destructive one. If we cannot know that we ever use words one way rather than another then we cannot know that we use them significantly at all. There is, therefore, no reason to accept the significance of the rule-sceptic's remarks. The sceptic might reply by claiming to dispute only past uses while clinging firmly to present use to do so. But that is exactly what is at issue: what is present use ? An indefinite number of possible interpretations are available. This is, therefore, no way out.

Rule-scepticism is disputed by Baker and Hacker [1986A] on a number of other grounds however. Firstly, the possibility of future divergence from past use might be accounted for in terms of misunderstanding the explanation of the rule. The fact of any agreement in past or present use, and surely such exists, is a fact which testifies to there being a shared understanding of a common rule. The sceptic apparently fears the future and its attendant potential for divergence. But such divergence is only possible if today's apparent understanding conceals misunderstandings which will generate, in Wright's [1980] phrase, "radical semantic variations". Crucially however, no explanation of a rule is ever identical with the rule, rather, explanation is an expression of the rule. The key assumption here is that following a rule must necessarily involve a prior interpretation of that rule. It is this

assumption which permits infinite, equi-viable interpretations at any given stage and thereby provides the sceptic with a foothold.

Baker and Hacker [1986A] are highly critical of many aspects of Kripke's [1984] account of rule-following. As is well known, Kripke's [1984] reading of remark 201 in the Investigations is one which focuses on the statement of rule-scepticism without taking seriously enough the remainder of that remark in which Wittgenstein describes an inherent "misunderstanding" in the formulation of the sceptical paradox. That misunderstanding consists however in positing an interpretation for every case of following a rule and inferring that the posited interpretation determines the meaning of the rule for its follower. But, Wittgenstein argues, "interpretations..do not determine meaning" and "any interpretation still hangs in the air along with what it interprets." [P.I. 198] There is, therefore, more to meaning than interpretation. As is the case with explanation, interpretation is not identical with the rule but is its expression.

It can be plausibly argued, I think, that to explain or interpret a rule is not completely constitutive of understanding the rule and though it contributes to such it by no means exhausts the meaning of the word governed by the rule. The effect of explanation depends upon reaction. It is precisely in appropriate reaction that understanding most importantly consists. In going on we manifest the fact that we know how to go on and exemplify understanding. The appropriate action and reaction need not be mediated by interpretation and explanation though these might be elicited on request. Hence:

"there is a way of grasping a rule which is not an interpretation but which is exhibited in what we call 'obeying the rule' and 'going against it' in actual cases"

(P.I., 201)

The rule is therefore expressed in corresponding explanation and interpretation. But if those explanations and interpretations diverge from the normal, where and only where meaning is clearly prescribed, then we do not attribute understanding but misunderstanding. In other words, to possess an interpretation, to explain, does not imply knowing how to go on. This alone does not entail the ability to follow the rule. If we do know how to go on then we react in a manner similar to those whose form of life we share. In *On Certainty*, remark 402, Wittgenstein quotes Goethe's proclamation that: "In the beginning was the deed". For Wittgenstein it is not justification or interpretation but action which begins the language-game. Action does not presuppose interpretation nor is justification in any sense presupposed. Action and reaction are at the bottom of language-games, beyond justification. Reaction may be responsively imitative, repetitive and instinctual but none of these characteristics require interpretation. In so acting and reacting, so doing, we partake in a custom and a culture which ultimately confers normativity and embodies the criterion of identity for following the rule. At this level, therefore, Baker and Hacker's [1986A] criterion for meaningfulness would seem to hold up rather well. That my explanation is acceptable to fellow language-users is a plausible criterion for my knowing what I mean.

But now it will no doubt be argued that rule-scepticism, or at least an appropriate analogue, could be reformulated to apply to the community. In this form the Sceptic's epistemological challenge demands a criterion for the identity of the rule at the community level. Baker and Hacker could of course argue here that the sceptic has just dug too deep. In seeking a criterion of identity at this level the sceptic is looking for

a justification of our practices. Here he has misapplied the concept of justification to something other than its correct extension. While the language-game prescribes what is to count as a justification within the boundaries of that game it does not make sense to ask for the justification of that which determines what is to count as a justification. It might be argued that as regards the sameness of the rule what the community counts as sameness is sameness. What the community prescribes, teaches, and commends as knowing how to go on, or, going on in the same way, just is knowing how to go on etc. This need not be a statistical rather than a normative point. The point is that both factors count here. The former is in harmony with the latter not in conflict. Of course, correct applications of "+2" or the word "red" are not settled by statistical survey but are inherited by the community or, perhaps, preserved within it. These are ultimately aspects of a culture and those who share that culture share certain common reactions and linguistic inheritances. The sameness of the rule is guaranteed in the first instance by our agreement in a language, in use, which is an agreement in practices. Secondly, it is assured by a more basic "agreement in judgements" and a sameness of instinctive reaction. Not surprisingly:

"the word "agreement" and the word "rule" are not unrelated to one another. They are cousins. If I teach anyone the use of the one word he learns the use of the other with it". (P.I., 224)

Most of these, and other, arguments are featured in Baker and Hacker's "Rules, Scepticism and Language" [1986A]. In employing them against Kripke I think these authors may already have gone further than is necessary. Kripke's account of rule-scepticism in Wittgenstein's "Philosophical Investigations" is given in Chapter 4 of his "Rules and

Private Language" [1984]. It is enough to refute that piece of commentary to re-examine Kripke's prime piece of textual evidence; remark 201, which I have already mentioned. Kripke's [1984] desire to attribute rule-scepticism to Wittgenstein far outruns the plausibility of the actual text for such a reading. In addition, that author might also be referred to Wittgenstein's attitude to scepticism in general which, I have argued, is consistently unsympathetic. Of course, regardless of whether or not Wittgenstein himself was a rule-sceptic, it remains an independent question how we are to deal with rule-scepticism and whether or not it is really acceptable to contend that what the community accepts as "going on in the same way" really is going on in the same way. Further, there is, I think, a price to be paid for taking the kind of approach we have been considering to rule-scepticism at the community level. That is, it just is not clear that there are no further meaningful questions about objectivity and the criterion of identity of use at the community level.

Baker and Hacker's [1986A] approach here is to challenge the picture of a cumulative history of use to which each new use is added. In the sense intended, it is argued, there simply are no new uses. The whole argument is held fallacious because applying the rule to something to which it has not previously been applied is not itself a new application of the rule but is a matter of going on in the same way, and doing that is not in any sense unprecedented. Explanations, it is held, terminate in series of instances plus a similarity rider. When the sceptic has been given this much he can be given no more, just because there is no more to give. To the accusation of irrationality Baker and Hacker [1986A] answer that the pattern of explanation elucidates a grammatical connection and this shows what is rational. There can be no such thing as justifying grammatical connections, asking why pain-behaviour is evidence for pain,

for example, because the act of querying grammatical connections is not the expression of doubt but of misunderstanding. That is to say, according to Baker and Hacker [1986A] the rule and its applications are, ultimately, inseparable. The use fixes standards of correctness and the practice of use is the rule. The relation between the rule and its applications, according to Baker and Hacker [1986A] is an "internal relation" and there can be no further questions, those authors argue, no possibility of doubting those internal relations. This point, I think, distills the disagreement I have been considering. For, in a sense, the sceptical arguments of Wright [1980] and Kripke [1984] recognise the internality of rule and applications, and establish that it is not clear that there is more to following a rule beyond applications. To the more sceptical authors however, it appears that, if all we have to go on is past use, then there is a serious indeterminacy about which rule we are in fact following, and therefore a further question as to whether we are actually ever following a rule. Of course, most contemporary authors are reluctant to accept a so-called 'rules as rails' conception, and are even more reluctant to take on board the obvious Platonist implications of such a view. Despite that fact, many such contemporary authors feel that, in following a rule, we keep faith with something, as it were, even if we are not, ultimately, clear about what it is we keep faith with. In contrast, Baker and Hacker hope not so much to answer such questions, but to prevent them from arising, to expose them as confusions, as pseudo-questions, pseudo-doubts. For Baker and Hacker, if we are to doubt at all we must just accept the internal relations. Doubting the internal relations is not, it is held, a kind of doubt but a kind of misunderstanding, a rejection of the very concepts we use. It is at this

point that the concept of a criterion is brought into play.

II CRITERIA AND INTERNAL RELATIONS

The term "criterion" is explained by Baker and Hacker [1986a] in two later sections: "Scepticism and Internal Relations" and "Internal Relations and Criteria". In those sections the authors first outline the notion of an internal relation as a non-inductive but defeasible relation embodied in grammar and exemplified by the paradigm of the relation of a desire to its fulfillment. The notion is further defined by the claims that a relation is internal if it is inconceivable that the relata be otherwise connected and impossible that any third entity should mediate between them. Being embodied in grammar internal relations must enter into the meanings of propositions. Indeed, Baker and Hacker [1986A] explicitly contend that if one "...disrupts an internal relation.." one thereby "robs the proposition of part of its meaning" (p.99). Internal relations, therefore, will be determinants of, but not constitutive of, the meaning of the relevant proposition. The concept of an internal relation is, fundamentally, a semantic concept. The criterial relation fits into this scenario as a proper subset of internal relations:

"...the relation between outward behaviour and an inner state which Wittgenstein [calls a] criterion is laid down in grammar...this internal relation does not rest on any third thing."
(p.110)

In relation to the paradigm case of an internal relation criterial relations possess a greater degree of complexity, just in virtue of the defeasibility of any criterially-based assertion. With reference again to the relation between inner state and outward behaviour, already acknowledged as criterial, it is held:

"...there is a point of similarity between the kind of internal relation involved in this case and the simpler

kind exemplified by the relation of a desire to its fulfillment.."
(p.110)

Criterial relations are thus represented as a higher-order, more complex kind of internal relation. Statements corresponding to internal relations are held to possess the characteristics of rules of grammar and therefore to be "tautologous" or "necessarily true" (p.112), but as regards criterial relations, it is held:

"..the statement 'if someone acts thus, he's in pain' is not a tautology....Criterial support is defeasible."
(p.110)

It is, I think, misleading, and strictly incorrect, to describe "criterial support" or "criteria" as defeasible and the point must be made that it is the criterially-based assertion, and only the assertion, which is defeasible. However, it is further held that to restrict internal relations to tautologous cases alone is to "narrow attention to a proper subset of statements that have the role of rules for use" (p.112). Baker and Hacker's [1986A] reasoning is, I think, that criteria are a subset both of the set of internal relations and of the set of rules for use. In this work then attention is shifted from criteria as meaners, as constitutive of meaning, to criteria (or the assertions which they support) as defeasible, internal relations determining meaning; limits or markers delineating content. Criteria remain fundamental to meaning in so far as any failure to construe the internal relations appropriate to a proposition correctly amounts to a distortion of the meaning of that proposition and therefore to a misunderstanding of it. This then is the most recent of Baker and Hacker's characterisations of Wittgenstein's use of "criterion".

In itself the position seems a strong, well-defined one in keeping with the general anti-theoretical standpoint as regards meaning developed

in "Language, Sense and Nonsense" [1986] but does it represent the later Wittgenstein's position ? It is an obvious initial strategy to seek confirmation of Baker and Hacker's [1986A] hypothesis in Wittgenstein's later work. Evidence for the hypothesis is, unfortunately, extremely scarce. Philosophical Investigations contains only one mention of the concept of an internal relation located in a discussion of 'seeing as' in Part II. Nothing of the relation between that notion and the notion of a criterion is explained in that passage. More to the point, no reference to any such discussions in the Philosophical Investigations is to be found in Baker and Hacker [1986A]. Instead reference is made to discussions in the Tractatus and the Philosophical Remarks. As is well known, much of Part One of the Investigations is devoted to severe criticism of many of the central ideas of the former work, including such concepts as simplicity, determinacy of sense, the relation of language to reality and so on. Proposition 4.014 of the Tractatus, for example, describes an internal relation of depiction, in the sense of shared logical form, between proposition and fact. My point is that even if this remark can be used as evidence for Baker and Hacker's [1986A] hypothesis those authors just have not presented the independent argumentative support to show how these ideas survive their own critique in Part One of the Investigations, and again, in what sense they are at work in the positive elements of that text.

Even the staunchest supporters of Wittgensteinian continuity would not accept that his position at the time of writing the Philosophical Remarks is identical either with that of the Tractatus or with that of the Philosophical Investigations. Again, my point is just that Baker and Hacker [1986A] do not offer any persuasive argument justifying the view

that what is often taken to be a brief verificationist phase in Wittgenstein's thought is actually more obviously consonant with the later Investigations. If Baker and Hacker [1986A] are correct they have not made clear why their readers should believe them to be.

There are at least two questions here which, I argue, the authors of "Rules, Scepticism, and Language" [1986A] have not lucidly distinguished. The first concerns the use made by Wittgenstein of the term "criterion" in his later work. The second and separate question concerns what can be made of the concept of a criterion in the contemporary philosophy of language. At this stage of the inquiry the latter must take precedence over the former for, although I have tried to cast doubt upon the plausibility of "Rules, Scepticism and Language" as an answer to the former question, we might readily concede the exegetical validity without having made any progress whatsoever in terms of the latter, more important, question. There is no room here for an argument from authority to the effect that because that is what Wittgenstein said it is therefore true. As regards Baker and Hacker's [1986A] conception of a criterion as a candidate for the correct notion which would be of value to the contemporary philosophy of language, rather than purely of interest to Wittgenstein scholars, there are a number of respects in which I wholeheartedly agree, but others where I have deep qualms.

Most importantly, it must be emphasised that the conception of a criterion developed in Baker and Hacker's [1986A] text, which conception is attributed to Wittgenstein, is itself, in a sense, a case not yet proven. That is to say, defeasibility, that characteristic which is alleged to distinguish criterial support from entailment, and is acknowledged by the authors to be both the most controversial and most problematic aspect of the notion of a criterion remains a deep problem.

Serious doubt has been cast on the very coherence of the concept and particularly on its connections with notions such as truth, knowledge and belief. Yet of this feature Baker and Hacker write:

"Philosophers have notoriously found difficulty in making sense of Wittgenstein's remarks about criteria. The crucial problem is to explain how an internal relation can be defeasible..Wittgenstein suggests that undefeated criterial support renders the proposition supported certain. How can this be ?"
(p.111)

In fact, I have already cast doubt on whether Wittgenstein suggested any such thing in Chapter One when I noted Wittgenstein's claim that the criteria do not put the sensation beyond doubt. Be that as it may, there is a further question beyond one of pedigree, namely, can sense be made of such a notion ? Baker and Hacker's [1986a] answer to their own question is, unfortunately, that:

"This is not an occasion for resolving these conundrums. Pro Tempore the merest sketch must suffice."
(p.111)

But if that is not the place to so pronounce then Baker and Hacker [1986A] attribute a position to Wittgenstein which they do not fully explain and further expect their readers to take on board a radical and controversial relation without tackling the deep problem which lies at the heart of this whole area. The argument in the sketch referred to runs as follows: the rule-sceptic infers from the fact that any criterially-supported assertion can be doubted that he can intelligibly doubt the criterial relations themselves. Further, the rule-sceptic fails to appreciate that to grasp the rule is to know what counts as following the rule. Now, the first of these points is clearly a good one. The rule-sceptic's argument seems plainly invalid. Equally, the second point appears reasonable. But the problem is this: do these points

satisfactorily explain how 'undefeated criterial support' makes a proposition certain ? Further, what is the sense of 'certain' being employed here ? If that term is used in a subjective sense, the asserter will feel certain, then we may concede the point. But if that term is being used in an objective sense then, I think, the following point must be made: in any case of an undefeated criterially-supported assertion there can be no guarantee that the assertion is true, that we have knowledge here, if the assertion is to be a genuinely defeasible one. Can we even make sense of the idea of defeasible knowledge ? This robs Baker and Hacker's argument of its glory, for it now looks as if there is a further question as to whether or not a criterially-supported assertion, understood as defeasible, is in fact a true assertion ? How do we ever get criterial knowledge ? No amount of equivocating with the term 'certain' will explain how undefeated criterial support adds up to knowledge, and, in this context, that must surely be the sceptic's point.

Given the very prominent role of the notion of an internal relation in Baker and Hacker's [1986A] account of rule-scepticism and the unanswered questions of defeasibility and knowledge and of defeasibility and meaning serious doubt must, in the last analysis, be cast on the effectiveness, firstly, of Baker and Hacker's [1986A] refutation of rule-scepticism and secondly, on the value of that notion of a criterion presented by those authors. Before turning to examine in detail the questions I have mentioned I propose to briefly outline Wright's [1982] first account of criteria.

III FIRST THOUGHTS

Wright begins his "Anti-Realist Semantics: The Role of Criteria"

[1982] by suggesting that a plausible account of the type of assertibility-conditions account of meaning which, for example, Dummett favours over truth-conditional accounts, might be given in terms of a notion at least relevantly similar to Wittgenstein's notion of a criterion. Criteria are initially characterised as defining a term, as being publicly accessible and as typically being multiple. Wright [1982] then goes on to explain the contrast between criteria and symptoms, as outlined in the Blue Book and then notes Baker and Hacker's [1974] claim that undefeated criterial support confers certainty on the supported proposition:

"Baker and Hacker take it...that the satisfaction of criteria confers certainty...if certainty is taken to apply to all and only statements which may not reasonably be doubted, then - since what may reasonably be asserted cannot simultaneously be reasonably doubted - the satisfaction of a statements criteria will...pre-empt any reasonable doubt about it; for it will now be part of the content of the statement that it may reasonably be asserted in those circumstances."
(p.244)

Interestingly, even at this early stage, in a footnote [6, p.244] Wright points out that

"...it is doubtful whether this is a sufficiently robust sense of 'certain'."

Wright [1982] suggests that the assertability-conditions theorist will distinguish between criteria and symptoms in order to give an account of misunderstanding in terms of misconceiving the criteria for a sentence. However, Wright [1982] notes that such an account does not yet make clear how misunderstanding could be manifest in behaviour. The difficulty is compounded by the endless possibilities of explaining apparently aberrant use in terms of "idiosyncratic" desires, beliefs etc. The problem, Wright [1982] concludes, is to provide an account of the

type of situation in which explanation in terms of misunderstanding is clearly the appropriate type of explanation. Further, if the criterialist wants to invoke criteria at this level he will have assumed that criteria are required to properly distinguish between explaining meaning and pointing to empirical facts. Wright [1982] goes on to stipulate a conception of criteria $\{D1..Dn\}$ for an assertion S such that questions as to the obtaining of the criteria are decidable and such that those criteria are S -predicative of S and therefore provide a non-circular account of the assertibility-conditions for S .

The criterialist now has two options: either there are cases of $\{D1..Dn\}$ supporting S which produce non-contingently true instances of that schema, criterial instances, or the description of the assertability conditions for S do not provide full understanding of S . If the latter view is correct ostension will be required to communicate understanding of the relevant conditions for S , hence the former view will deny the possibility of ostension. If we take the ostensive alternative, Wright [1982] argues, we might ask what grounds the community's belief that the expression really is warranted by conditions of the type ostensively explained to us. The interesting response will be one of puzzlement and inability to explain hinting at a conventional relationship, rather than a response which tries to adduce empirical evidence. The first response, Wright argues, describes conventional support and characterises the notion of *criteria:

"..the connection..is one of conventional support if and only if we would consider it acceptable in certain circumstances to assert S on the basis of such a state of affairs ; and we would not require, in order for someone to be credited with a full understanding both of S of what state of affairs had been ostensively explained, that he know what it would be empirically whether occurrences of that type of state of affairs really did provide reason for believing S and find that they did not."
(p.252)

"...a particular state of affairs is a *criterion..if and only if it affords conventional support in the sense just explained. Truth-conditions ..are a sub-class of *criteria"
(p.252)

If meaning is to be explained here, Wright [1982] argues, then recourse cannot be made only to symptoms, viewed as giving defeasible empirical support. For the model itself cannot make clear those defeasibility-conditions. Hence the assertibility-conditions theorist must have recourse either to criteria or to *criteria. Wright goes on to note the obvious problem of criterionless cases for the criteria-inclined assertion-condition theorist:

"Our question now is: will an assertability-conditions theorist who is ready to grant that a particular class of verification-transcendent statements possess no criteria be committed to their possession of *criteria instead ?"
(p.252)

The difficulty here, as Wright points out, is that:

"Failing recourse to criteria or *criteria..the theorist lands himself in the incoherent position of having to grant that there is an empirical question about the adequacy of any particular explanation of the use of S into which no effective..investigation can be conducted."
(p.253)

Wright, of course is interested, primarily, in a more general, even global, notion suitable for use in a theory of meaning based on assertibility-conditions, as opposed to truth-conditions. From such a standpoint then either knowledge of criteria, *criteria, or both will be prerequisite to adequate explanation. Wright goes on to explore the plausibility of such an approach with reference to the test-case of statements about the past.

The existence of standards of correctness, it is held, and the ability of the community to produce a mastery of such language in the linguistically mature must at least protest against, if not refute, the claim that there

can be no criteria for past tense statements, but the issue is not fully prosecuted here. Wright [1982] now follows Baker and Hacker in concluding that a criterial theory could not be a global theory of meaning:

"If, as in Hacker and Baker, the criterial relation is regarded as holding between statements, then an anti-realist who held that the meaning of every statement is determined by its association with criteria would commit himself to an impossible model of how an understanding of any statement could be acquired."
(p.258)

The impossibility, as Wright notes in a footnote (14, p.259) arises from the fact that the criteria for each criterially-explained statement must themselves be explained in terms of criteria which again require criterial explanation, and so on, ad infinitum.

Therefore, it is argued, the criterialist assertibility-conditions theorist must recognize a base class of statements whose meanings are not criterially given, if he is to escape from the hermeneutic circle. Hence, Wright [1982] subsequently explores the question of the extent to which the anti-realist can accept a truth-conditional analysis, at least, in order to determine the nature of that base class. Among the possible candidates for a class of effectively decidable propositions observational statements appear the most initially plausible. But, Wright [1982] notes that the verification of any such statement presupposes a notion of spatio-temporal positioning and the cognitive map which allows judgements about that to be made is itself a defeasible and revisable empirical theory which therefore cannot provide the kind of conclusiveness required by the notion of verification; particularly given the possibilities of illusion. If we have just chosen the wrong candidate, it is held, then the right one must be indexed to the here and

now and be devoid of proper names, definite descriptions and sortals which presuppose futures or pasts. Worse, the possibility of an abnormal position or state of the observer drags defeasibility back into the picture even in the retreat to "seemings" and we are pushed further back to a degree of privacy which is nothing other than completely discredited [ie. Chapter Two, above]. Thus, it is held, there is no special problem about undecidable statements, for defeasibility is simply all-pervasive.

If there is no plausible base-class, Wright argues, then two options are available:

"The first course would be to look to a base class of statements whose meaning was determined by their association with *criteria. But of the realization of these *criteria there would be provision for no other description save by the use of the very statements whose assertion they non-contingently supported."
(p.265)

"The second alternative is to break with the assertibility-conditions conception of meaning and to look to a base class the correct use of whose members, while something on which any opinion is essentially defeasible, is thought of as founded..in human reaction alone..an anti-realist who takes it [this option] will have little justifiable complaint if he is regarded as a neo-idealist."
(p.266)

In conclusion, I do not propose to critically assess Wright's [1982] remarks here, although it will be clear that Wright is dealing with a notion of a degree of generality the very idea of which I have already been highly critical. I will however attempt to assess that authors own second thoughts [1984]. In each of the conceptions of criteria I have outlined in this chapter the notion of defeasibility has an undeniably definitive role and yet so does the desire of each author to preserve a notion of certainty, of truth or of knowledge. The question of the compatibility of these two considerations forms the basis of Wright's [1984] later critique and it is to the question of that uneasy marriage of epistemological notions that I turn in the next Chapter.

CHAPTER SIX

INTRODUCTION

Crispin Wright's "Second Thoughts About Criteria", published in 1984, has presented contemporary philosophy of language with the most sceptical, provocative, challenging and demanding discussion of the concept of a criterion. In that paper the concept of a criterion is described in terms of the following five cardinal features:

"...that recognition of satisfaction of criteria for P can confer sceptic-proof knowledge that P; that P's criteria determine necessarily good evidence for P, and thereby fix its content; that the criteria for P will typically be multiple; that satisfaction of a criterion for P will always be a public matter; and that to know of the satisfaction of criteria for P is always consistent with having, or discovering, further information whose effect is that the claim that P is not justified after all."
(p.383)

It is then argued that:

"...it is seriously unclear whether...the fifth feature, defeasibility, can be made to harmonise either with the first - the Knowledge feature or with the second - the Meaning feature."
(p.383)

Further:

"But criteria would not be interestingly different from public truth-conditions if defeasibility was waived; and jettison of the meaning feature would threaten the knowledge feature as well. While to forgo the knowledge feature would be to deprive the notion of what most advocates have seen as its principal point."
(p.383-4)

In fact the paper unfolds a range of problems which cast doubt upon the coherence of the concept of a criterion so described and raises the question of the relationship between the concept so characterised and Wittgenstein's use of the term "criterion". In what follows I will discuss each of the problems outlined and will assess the allegedly

Wittgensteinian pedigree of each of the features of that conception of a criterion.

I DEFEASIBILITY AND KNOWLEDGE.

The first question raised is that of the possibility of reconciling defeasibility with knowledge. The point is formulated as a "simple objection":

"If I claim to know that P I will be understood to be claiming that my belief that P is guaranteed correct; so I must have a conclusive basis for that belief if my claim is to be true. But knowledge that criteria for P are satisfied is to be consistent-by defeasibility-with my obtaining further information as a result of which I no longer have a basis for the belief that P at all. Information that is genuinely conclusive for a certain belief, however, cannot lose that status as a result of being added to; so knowledge that criteria for P are satisfied cannot ever amount to knowledge that P-the two states have different essential characteristics."
(p.384)

The same point is emphasised again on page 386:

"...for to stress, recognition of satisfaction of criteria for P cannot, consistently with their defeasibility, be held to constitute knowledge that P without contravention of the truism that knowledge entails the truth of what is known."

Wright [1984] explicitly acknowledges his debt to John McDowell's "Criteria, Defeasibility and Knowledge" [1982] for the simple objection ['3', p.384, '3'-notes, [6]-references]. McDowell's [1982] doubt about the viability of criteria is expressed in the first section of his paper and a brief quote makes clear the same epistemological worry motivating Wright's "Second Thoughts":

"...the criterial view does envisage ascribing knowledge on the strength of something compatible with the falsity of what is supposedly known. And it is a serious question whether we can understand how it can be knowledge that is properly so ascribed."

(p.458)

What the point comes to is that while knowing that P entails that P is true, which in turn entails the fact that P, having criterial reasons to claim to know that P cannot entail that P is true or that P is in fact the case; if the criterial assertion is to remain a defeasible one. What is lost in conceding the point is precisely the first feature: "that recognition of satisfaction of criteria for P can confer sceptic-proof knowledge that P."

Suppose then that we do not concede the point. What must be provided to answer the challenge of the simple objection? The key to the form any answer must take is given in the following phrase from the simple objection: "...obtaining further information as a result of which I no longer have a basis for the belief that P at all". In other words, underlying the simple objection is the formal point that the logic of the concept of truth is monotonous:

1. $K_{t_0} p \rightarrow T_{t_0} ('p') \rightarrow (\forall t) T_t ('p')$, where ' t_0 ' is intended as some fixed "now" or present moment.

If I know that P at some given moment then not only must 'P' be true at that moment but 'P' will always be true and, indeed, 'P' was always true. Therefore, it is a notion of timeless truth which the formal claim of monotonicity of truth brings into play here. Of course, there is the independant claim, specifically about knowledge, that:

2. $K_{t_0} p \rightarrow (\forall t)((\text{if } t \text{ is later than or equal to } t_0) \rightarrow Kp)$.

If P is known at a given moment then P is known at that moment and at all times later than that moment. While this might be said to hold for objective knowledge, in a Popperian sense, it is far less plausible as regards more mundane knowings, or any account of knowledge involving reference to belief states. One obvious objection would be, simply, that

we might forget. In conclusion, were it not for the claim about the monotonicity of truth formulated above, the criticism of (criterial) knowledge involved in the simple objection would be less than perspicuous. Consequently, what is required to answer the challenge and to save the first feature of a criterion is a non-monotonic logic in which clear sense can be given to the notion of 'defeasible knowledge'. That is, a logic must be developed which validates, has as a theorem, the denial of the thinning condition familiar from classical logic, ie. a logic whereby the addition of information 'Z' to the left hand side of the formula 'X entails Y' can bring it about that Y is now not entailed at all. In effect, I argue, the first (simple) objection already involves the second of Wright's [1984] criticisms and in addressing myself to the underlying formal point I hope to say something of relevance to both. However, in my own treatment of these questions I follow Wright's [1984] order of exposition.

Recent attempts to formalize just such a notion of non-monotonicity have certainly been made. What is less certain is whether any of them can deliver the goods. In order to answer that question the proposed logics must be critically assessed. In Chapter Five of his "Logics for Artificial Intelligence" [1984] Raymond Turner defines non-monotonic logic in the following way:

"Non-monotonic logical systems are logics in which the introduction of new information (axioms) can invalidate old theorems"
(p.61)

The definition is illustrated by an example borrowed from Dov Gabbay [1981]. The situation described is as follows. A tour operator offers flights from New York to Paris and Stuttgart. At 12:00 hours Stuttgart discovers that due to a terrorist occupation of Paris airport all Paris

flights will be re-routed. Stuttgart also knows that the New York flight took off and presumes that it did so on schedule. Stuttgart then assumes their flight will be re-routed to London and infers that its passengers will arrive tonight from London. This conclusion is referred to as 'D'. At 14:00 hours however Stuttgart learns that take off had been delayed by two and a half hours. It is then assumed that the pilot would have returned to New York and it is concluded that '-D', the passengers will not arrive tonight.

If non-monotonic logic is to provide the basis of a reply to the simple objection then this example ought to be similar to the case of criterial reasoning. It might be argued that the case parallels the criterial one quite well. In each case an expanded state of information leads to a conclusion which contradicts a conclusion made earlier on the basis of a thinner state of information. The problem as regards knowledge is that if we now know that -D at time t_2 then, by monotonicity, we did not know that D at time t_1 . Otherwise it looks as if we can deduce a contradiction and refute 'non-monotonic knowledge' by Reductio Ad Absurdum.

Turner [1984] considers a number of attempts to formalize, in a coherent way, the condition of non-monotonicity. The first system is that of McDermott and Doyle [1980] and its subsequent elaboration in McDermott [1982]. McDermott proceeds by introducing a consistency operator "M" to a first-order modal language, L_n . A possible worlds semantics is given proceeding from the notion of a modal frame "M" for L_n as the quadruple [W, D, R, F] such that:

- "(i) W is a non-empty set (of possible worlds).
- (ii) D is a non-empty set.
- (iii) R is a binary relation of accessibility between elements of W.

(iv) F is a function which assigns to each pair, consisting of an n -place function symbol and an element of W , a function from D^n to D and to each pair, consisting of an n -place relation symbol and an element of W , an element of 2 to the D^n .
(p.62)

The relation: " M semantically entails g, w, A " read as: assignment function g satisfies the wff A at world w in frame M , is then recursively defined.

McDermott [1982] and Turner [1984] argue that different properties of the accessibility relation ' R ' generate distinct modal systems. When R is reflexive modal logic T is obtained. If R is reflexive and transitive $S4$ is obtained. Finally, if R is the equivalence relation $S5$ is obtained. The semantic basis of McDermott's account of non-monotonic inference is therefore a monotonic one based on the notions of semantic entailment in T , $S4$ and $S5$ respectively. Turner [1984] goes on to identify a rule of inference, possibilitation, as the heart of McDermott's [1982] account of the non-monotonic fragment of the logic: "'Can't infer $-A$ ', then MA ." ie. if $-A$ is not provable then A is consistent. (There is in fact a misprint of " A " for " $-A$ " in Turner's text.) As Turner [1984] notes:

"..the notion of inference in the phrase "can't infer" is not to be identified with any of the relations, semantic entailment in T , semantic entailment in $S4$, semantic entailment in $S5$, but is to be obtained from such a relation by ..bootstrapping..; we require a relation of entailment..which satisfies: If $-A$ is not entailed then MA is entailed."
(p.63)

Here I can do no better than to cite the concise account of the formal manoeuvrings given by Turner [1984]:

"We begin with an enumeration ($A_i, i = 1, 2, 3, \dots$) of the wff's of L_n . Let S be some set of premises and "semantic entailment" some underlying notion of entailment. Then put:

$S_0 = S$

and

L_n - if for some B in L_n MB is a member of S_i and S_i entails $\neg B$

$S_{i+1} = \begin{matrix} S_i \cup (MB_i) & \text{if } S_i \cup (B_i) \text{ is consistent} \\ S_i & \text{otherwise.} \end{matrix}$

We then put $S_{\aleph} = \bigcup S_i$. Obviously, S_{\aleph} may depend upon the particular enumeration selected. As McDermott & Doyle point out that if $S = (MC \rightarrow \neg D, MD \rightarrow \neg C)$ then either MD or MC will be selected (but not both) according as C or D is encountered first in the enumeration. To circumvent this problem McDermott and Doyle make their relation of non-monotonic deducibility independent of any particular enumeration by defining: S non-monotonically entails A iff. S_{\aleph} semantically entails A , for every enumeration of L_n . As can easily be seen non-monotonic entailment satisfies the so-called rule of possibilitation." (p.63)

This then is a formal theory of non-monotonic inference. But before we can ask how this might help in the criterial context it must be noted that it is not yet a wholly satisfactory theory. Turner [1984] notes that if a formula A is entailed non-monotonically in S_5 then A is semantically entailed in S_5 anyway. In other words, non-monotonic S_5 demonstrably collapses into monotonic S_5 . While this shows that non-monotonic S_5 is consistent it equally demonstrates the redundancy of all non-monotonic reasoning in that context; a highly undesirable consequence. Conversely, the fact that non-monotonic T and S_4 do not collapse into their monotonic counterparts shows that non-monotonic reasoning in these contexts is not redundant but does leave the question of the consistency of these systems hanging in the air. In fact Turner points out that for the non-monotonic systems the theory, $LMC \rightarrow \neg C$ which is consistent in monotonic S_4 and T is inconsistent in the non-monotonic systems. There are therefore concrete grounds to doubt the consistency of these systems.

Further, McDermott and Doyle omit any semantics for the consistency operator "M" from their original [1980] presentation. One interesting implication of "M" lies in the consequences it has as regards the notion of possible worlds in the system. Clearly, if "MA" is to be understood as the consistency of A then possible worlds cannot be understood in the ordinary way as determinate states of affairs (or information). Rather, the picture is of a world evolving as a developing state of information such that, initially, we may have neither "A" nor "-A" and, subsequently, as the state of information develops, it becomes the case that A. The law of excluded middle cannot hold for such worlds which simply are not totally decided or determinate at any given stage. If the worlds were determinate then it would be difficult to see how any two worlds could be anything other than strictly inconsistent.

In short, it just is not clear that the McDermott/Doyle system, in its present form, provides a coherent picture or a workable logical system which can stand on its own legs. Therefore, there can be no real hope of using this formal theory, just as it is, as a basis for a reply to the simple objection from monotonicity in the case of criteria.

Dov Gabbay [1982] provides an alternative approach to the problem from an intuitionistic basis which exploits the notion of (increasingly complete) states of information over moments of time. Again, reading 'M' as 'plausible' or 'consistent' I quote Turner's definition:

"A Gabbay Structure, $T = [T, \text{"less than"}]$ where T is a non-empty set (moments of time) and "less than" is a reflexive and transitive relation on T , that is, for each t, t', t'' in T we have t "less than" t and (t "less than" t' & t' "less than" t'') implies t "less than" t'')." (p.65)

Gabbay also exploits a function h which assigns the values 1 or 0 to each atomic formula at a moment of time and the function is applied to

every wff by recursive clauses. In addition h is subject to the following stipulation: if h assigns the truth-value surrogate '1' to a formula at t , h must also assign 1 to that formula at t' .

Turner notes that Gabbay's account of implication derives from Kripke's semantics for intuitionistic logic and that:

"The operator M is interpreted as true of a wff, at a moment in time, if it is true in some possible continuation of that time."

It is also noted that $\neg A$ and MA are mutually exclusive in the sense that both cannot hold at the same moment in time. Gabbay defines a notion of provability, A proves B , as, if function h assigns 1 to A at some moment then h so assigns to B . Gabbay claims three advantages for his system over the McDermott / Doyle system. Firstly, the rule $\neg MC / \neg C$ which was invalid in the early system is validated in Gabbay's system. The same is also true of the rule $M(A \& B) / MA$. Finally, $(MC, \neg C)$ is inconsistent under the Gabbay regime but is not so for McDermott and Doyle.

Turner goes on to explain Gabbay's notion of non-monotonic provability by reference to Reiter's [1980] default reasoning, where a "default" is of the form:

" $A(X) : MB_1(X), \dots, MB_m(X) / C(X)$ "

where $A(X)$, $B_1(X) \dots MB_m(X)$, $C(X)$ are wff whose free variables are those of $x = x_1 \dots x_n$. $A(X)$ is called the prerequisite of the default and $C(X)$ the consequent. An example of default reasoning is the following:

$BIRD(X) : MFLY(X) / FLY(X)$ we are to interpret this.. as 'if x is a bird and it is consistent to assume that x can fly then infer that x can fly'. The crucial observation here, of course, is that some birds cannot fly.

Gabbay's rule of non-monotonic inference is just the transitive closure of Reiter's default rule."

(p.67-8)

While Gabbay's [1982] system might represent an improvement on the McDermott / Doyle system in the respects claimed, actually applying Gabbay's system to criterially-based language remains problematic. Of course, that system was not originally addressed to the problems of criteria. Despite this fact, criterial interpretations of the system can be made. These interpretations are mine and not Gabbay's. One interpretation might be illustrated by reference to the following principle which holds on Gabbay's account: $MA \vee \neg A$. Presumably, "MA" ought to be read along the lines of "A is not at present criterially ruled out". That reading would however raise questions about what it is to rule out an assertion criterially. One obvious response, though not the only one, would be that A is ruled out criterially just in case A is defeated. On this view a criterial negation will just be a defeated assertion. Now, $MA \vee \neg A$ could then be taken to say that: 'either A has not been defeated or A has been defeated', that is, either A is defeated or it is not. If the obtaining or not of defeaters is taken to be a decidable matter and non-monotonicity does not occur at the level of criteria and defeaters, then it is to be expected that at that level the law of excluded middle will hold. In other words, at that level on the suggested construal, $MA \vee \neg A$ would act as a surrogate for the ordinary law of excluded middle. If A is defeated then it is ruled out and we may (criterially) assert its negation. If however A is not defeated then, while we may have no basis actually to assert A, A is clearly not ruled out. Therefore, we may conclude that MA, A is consistent.

Here, of course, we have presupposed that we can simply read " $\neg A$ " as meaning that "A has been defeated" which would mean that the truth of negations, and therefore falsity, is in fact monotonous while the truth

of unnegated sentences is supposed not to be. However, the alternative to accepting such an asymmetry might itself lead to far less plausible claims. An obvious strategy to resolve the asymmetry might, for example, involve appeal to a more complex theory involving positive criteria for negated sentences ie. criterially-based claims about the falsity of negated sentences; but it is not clear that any such theory could remain faithful to the conception of criteria under discussion. The first job of a criterion is surely positive, as a basis for the positive assertion of something. I would want to deny that we ever have negative criteria, criteria for the absence of something. In a case of, say, pretence, I would not want to argue that we have a criterion for someone's not being in pain. That someone is not in pain surely follows from the internal conceptual relations which prevail here but recognition of the satisfaction of pretence-criteria primarily provides a basis for the positive assertion that someone is pretending and not for the absence of anything. I would want to argue that we do not assert $\neg A$ criterially, rather we withdraw the assertion of A because of some defeater D . That is, it seems far more plausible to argue that rather than there being criteria for someone's not being, say, happy that someone was not happy could follow from their being, say, in pain, for which there existed direct criterial support. In other words negations of this type follow not from 'criteria for negations' but from internal meaning-relations. Formally then, there could be no reading for $C \rightarrow \neg A$ other than $(\exists DA) DA \in W$, where ' DA ' is a defeater of A and W a specified context. On this account we might explain $C \rightarrow (B \rightarrow \neg A)$ as: $\neg(C \rightarrow B) \vee (\exists DA) DA \in W$. Again, the rider that the first disjunct is not assertable in isolation must be stipulated.

It must also be emphasised that it just is not obvious that any defeater is itself a criterion for something else. Although that possibility need not be ruled out completely, the claim that every defeater is always a criterion for something else seems an unrealistically strong claim. It must be more realistic, I argue, to claim that negations are introduced derivatively from conceptual interconnections formally reflected, perhaps, in meaning-postulates while the primary purpose of criteria remains positive. Thus, read purely criterially, I would argue, it is not clear that sense can be made of the claim that $B \rightarrow \neg A$. Therefore, rejecting the asymmetry outlined in my first account seems to force upon the criterialist a story about negation which might be the simplest way to avoid asymmetry, but which may also involve sacrificing some very plausible intuitions about criteria and negation. I return to this subject again later in the present section. At this stage however, as regards Gabbay's system, I argue that although it does represent a substantial improvement it still leaves open which type of criterial interpretation the formal system admits and thus it remains an open question how that account might form the basis of a non-monotonic theory of criteria.

Finally, with the notion of a partial model, intended to represent an incomplete state of information, Turner [1984] presents his own account of non-monotonic reasoning:

"A partial model for L is a structure $M = [D, F]$ where D is a non-empty set and F is a function which assigns to each n -place relation symbol C^n (for each n greater than 0) of L an n -place function C^n from D^n to $\{t, f, u\}$ "
(p.69)

Turner proceeds to formally characterize the notion of a plausible extension of a partial model understood as an increasingly complete state of information, as follows:

"Let M, M' be partial models of L . We shall say M' is an extension of M just in case for each n -place relation C , of L (each n greater than 0) $C^{M'}$ is an extension of C^M (considered as functions from D^n to $\{t, f, u\}$). We write M "less than" M' to signify that M' is an extension of M if M' is a plausible extension of M then M' is an extension of M ."
(p.70)

The relation is held to be reflexive and transitive. A semantics is then set up, based upon a three-valued system due to Kleene.

In the last analysis however Turner concedes that:

"Of course, just as in the case of the accessibility relation between possible worlds in modal logic, we cannot spell out exactly what this means. The notion will remain somewhat vague".
(p.70)

This, I think, is a poignant and acute observation which must, at the end of the day, cast a shadow of doubt on the ability of such formal systems to offer a solid basis for a reply to Prof. Wright's [1984] 'simple objection'.

It is also conceded that the non-monotonic system is much less "well behaved" than its monotonic counterpart. In other words, despite the fact that both Gabbay and Turner have developed systems which claim to be improvements on the original McDermott / Doyle system, the same types of difficulty of both the formalism and the underlying intuitions re-emerge, albeit in a different form. In short, the current, highly programmatic, state of the semantic theory of non-monotonic inference does not presently provide any sort of basis for a general counter to the appeal to the monotonicity of the concept of knowledge. Neither is it at all clear how any existing system could be read criterially, in particular. Despite this, if one is to uphold the kind of conception of criteria which Wright [1984] outlines, a solution to the problem of non-monotonicity which is presupposed must be developed. Failing that, the

first of the cardinal features must therefore be jettisoned as, currently, indefensible.

A better approach to the specific problems raised by criterially-based language might be to proceed from the concept of a "context"; where a context is understood to be composed of true or false sentences describing criteria and defeaters. The truth-value of P, the criterially-based assertion, would be computed on the basis of the assignment of truth-values to the criteria and the defeaters which constitute the context. The theory would have to be a recursive one at this point including both ordinary connectives and a criterial operator. P, the criterially-based assertion, will be true in a context when descriptions of the criteria in that context are true and descriptions of defeaters are false. P will be false when the descriptions of the defeaters are true. A context might be understood as analogous to the notion of a world employed in much contemporary modal semantic theory and at that level the logic of each context would be classical and therefore bivalent.

Although this approach exploits the conceptual apparatus which modern 'possible worlds semantics' provides I do not take a realist position in my understanding of the formal theory. Rather, I hope to give the basic formal framework a primarily epistemic interpretation in terms of notions such as those of belief-sets or assertability. This understanding will cut free the metaphysical and ontological baggage which comes with realist readings and will permit the modelling of criterial assertion without taking the ontologically unnecessary on board. The semantics is presented only as an illustration, or formal explanation, and makes no claim to exist in the brains of criterial reasoners in any Chomskian sense [above Chapter 4]. Despite my

reservations about those claims however it would be absurd to argue that formal theory qua formal theory is philosophically misleading just in itself. Rather it is the status of the formal theory, the ontological and metaphysical claims made on its behalf which might be objectionable. No such claim is made here. The status of the theory is that of an illustration.

What can be said here will only amount to a sketch for a semantic theory but the basic picture is, I argue, fundamentally correct.

Further, it should be noted that I am indebted to D. Lewis [1973] whose work "Counterfactuals" has been a source of inspiration for many of my own ideas though, of course, none of my errors can be blamed on that or any other author. As stated the theory will be a modal one which will require a recursive theory specifying the combinations of criteria and defeaters indicated informally above, together with a recursive theory for P , the criterially-based assertion, in complexes somewhat as follows:

1. Base Clause: $C \rightarrow P$

In every case P is understood as prefixed by " $C \rightarrow$ " representing criterial relative necessity, as assertability on that basis is the notion in play throughout. Only those contexts where C is true, where the criteria actually hold, are considered. I do not consider contexts in which the criteria do not hold. $C \rightarrow P$ will be true iff there is no world in which P is defeated which is less far fetched than any in which it is undefeated. Otherwise $C \rightarrow P$ will be false. The notion of far fetchedness involved presupposes a formal relation of closeness or similarity of possible contexts to the actual context. I characterise that relation as "...being at least as similar to the actual as...". This relation will be transitive because when a context A is at least as similar to the actual

as a context B and context B is at least as similar to the actual as context C then context A will be at least as similar to the actual as context C. It will also clearly be reflexive in that every context will be at least as similar to the actual as itself. The relation will not be a symmetrical one however for if we consider two contexts such that the latter is more far fetched than the former then while the former is at least as close to the actual as the latter the converse obviously will not hold. Further, because connectivity does not hold here, there is no guarantee that it will always be possible to compare any two contexts for similarity to the actual. Intuitively then a criterial necessity will be one which holds over all the equally plausible nearest contexts.

2. Negation:

It will be clear from the discussion of Gabbay above that I will not claim a sensible reading for any form of criterial denial ie. 1. $C \rightarrow P$ will not have a reading. Instead I will deal with what must be an implication of 1: 2. $\neg(C \rightarrow P)$, read as, even if C holds it does not give us an inference-ticket to P, because there is a defeater showing in the ordering. While 2 can be seen as an implication of 1 it is clearly a weaker claim than 1. 1 and 2 are not therefore equivalent. Negation is primarily external, both because my interest here is in assertability and non-assertability and in virtue of the fact that there is no strictly criterial denial, because there are no negative criteria.

3. Conjunction and Disjunction:

- (i) $C \rightarrow (P \ \& \ Q)$ iff. $C \rightarrow P$ and $C \rightarrow Q$.
- (ii) $C \rightarrow (P \vee Q)$ iff. $C \rightarrow P$, $C \rightarrow Q$ or both.

4. Implication:

(i) $C \rightarrow (P \rightarrow Q)$ iff. $(C \rightarrow P) \rightarrow (C \rightarrow Q)$.

Further,

(ii) $C \rightarrow (-P \vee Q)$ iff. $(C \rightarrow -P) \vee (C \rightarrow Q)$.

So, from $(-P \vee Q)$ it follows that $((\exists Dp) Dp \in W) \vee (C \rightarrow Q)$.

However, I do not take the first disjunct which indicates that there exists a defeater of P as being equivalent to $-(C \rightarrow P)$. Again, while $(P \rightarrow Q)$ follows from $(-P \vee Q)$, the latter will not follow from the former. Further, if $(C \rightarrow -P)$ then $-(C \rightarrow P)$, by definition. Hence, the antecedent of (i) is false, and the (material) conditional must therefore be true.

I take $C \rightarrow -(P \vee Q)$ as equivalent to $(\exists Dp) Dp \in W$ and I take $C \rightarrow (-P \& -Q)$ as equivalent to $(C \rightarrow -P) \& (C \rightarrow -Q)$, ie. $(\exists Dp) Dp \in W$ and $(\exists Dq) Dq \in W$. Further, $C \rightarrow (P \vee Q)$ is taken as equivalent to $C \rightarrow (-P \rightarrow Q)$ from which it may be inferred that: $(C \rightarrow P) \vee -((\exists Dq) Dq \in W) \vee -((\exists Dp) Dp \in W) \vee (C \rightarrow Q)$. Finally, I think it clear that such an account can give a place to a criterial surrogate for the law of excluded middle, $C \rightarrow (P \vee -P)$, read as $((C \rightarrow P) \vee ((\exists Dp) Dp \in W)) \vee -((\exists Dp) Dp \in W) \vee -(\exists D-p) D-p \in W$. Therefore, while excluded middle does hold, I do not expect bivalence to hold as regards criteria; for while an assertion may not be defeated there may not be a basis to assert it. Further, it seems clear that dissection will not hold. Finally, the metalanguage would appear to be a classical one in that we can infer from, $(C \rightarrow B) \rightarrow (\exists Da) Da \in W$, that, $-(C \rightarrow B) \vee (\exists Da) Da \in W$. Of course, this can amount to no more than a sketch of a formal system although, I argue, it must indicate the right way to proceed.

As regards criterial implication: when $B \rightarrow A$ holds, the nearest worlds are such that whenever B holds, A holds. If now however some C turns up such that when both B and C hold A does not, ie. C is a defeater

of A, the worlds will re-order such that the nearest are now those where B and C hold but A does not. It is important to note that " $B \rightarrow A$ " still holds but if " $(B \ \& \ C) \rightarrow A$ " holds then B is no longer sufficient for A. The logic of criterion cannot therefore be represented in only one set of worlds but requires the transformations as well. It is in terms of the transformational re-ordering that I would hope to find the basis of an explanation of the non-monotonic features of criterial language explaining the transformations cognitively, in terms of change in one's picture of the world-ordering. The transformational moves might, it is hoped, be a function but, of course, defining it is a deep problem, and my account remains no more than a sketch for a helpful illustration. Although the suggested approach does seem to be an improvement it cannot in its present form provide a solution to the problem raised by Wright's [1984] simple objection.

That problem is only a problem for the concept of a criterion as characterized by the five cardinal features, however, and is not any kind of problem for an alternative conception, which does not take the point of the concept to consist in the acquisition of sceptic-proof knowledge; if indeed that is a real possibility. Before outlining my own conception of a criterion I will consider the question of the (allegedly Wittgensteinean) pedigree of the Knowledge feature. The answer to this latter question will, I argue, shed light on the former.

I have already argued that Wittgenstein did not use the term "criterion" to refute scepticism in the Investigations [above Chapter 3], and this in itself would be surprising if that author really had held that criteria could provide sceptic-proof knowledge. Further, I have also noted [Chapter 1] Wittgenstein's claim, in the Remarks on the Philosophy

of Psychology Vol. 1, that recognition of the satisfaction of criteria does not put the existence of the sensation beyond doubt. Again this would be a surprising remark if Wittgenstein had held that recognising the satisfaction of criteria could confer sceptic-proof knowledge. More to the point is the reason which is given for the fact that the existence of the sensation is not put beyond doubt, namely, that there are no such criteria. Of course, it could be that Wittgenstein did adhere to a conception of criteria which incorporated the knowledge-feature, but thought it inapplicable in the case of sensations. Even if that were the case however, the point would still hold that the concept of a criterion nonetheless was not introduced as a way of refuting scepticism across the board.

In Chapter 1 I tried to explain Wittgenstein's role with respect to contemporary characterizations of criteria, and argued there that many features of contemporary theories are simply not present in Wittgenstein's later work but are attributed to him only by certain commentators. This is a case in point. But there is more to be said. Regardless of pedigree the first feature is plainly indefensible as it stands. What might be both defensible and of value against scepticism is a distinction with which I replace the first feature I have jettisoned. That distinction is one between the imaginability of the negation of P and the existence of a ground to doubt that P. This distinction is integral to any reasonable conception of a criterion regardless of pedigree. I will therefore first make clear the nature of the distinction and its point in the context of scepticism before examining its breeding, as it were.

It will already be clear that in relinquishing the knowledge feature we are left with, at strongest, a belief feature. Rather than discussing truth and knowledge therefore I will discuss justification and belief. I am aware that even the coherence of criterial versions of these notions is challenged in Wright [1984], but as a first step towards salvaging a coherent notion of a criterion a few preliminary remarks are necessary.

From the standpoint of a criterion as a basis for belief, as a basis for claims to know, understood defeasibly, it can be seen that there is an ambiguity in the formulation of the Knowledge Feature: "to know of satisfaction of criteria for P is always consistent with having further information whose effect is that the claim that P is not justified after all." To be clear, the point is not that after all P was never justified, but that after all P might not be true. While truth may be analytic of the concept of knowledge it clearly is not analytic of the concept of justification.

If what we are talking about here is criteria as reasons to believe, or defeasible warrants to assert, then the following, rather Popperian, point might be argued against the need for any such notion. It is often held that speakers frequently just say without recourse to grounds or warrants of any sort. No doubt it would be blind to ignore this fact, but here it is worth stressing the importance of reason-giving to speakers. In the epistemological context it seems just as blind, just as pointless, to deny that speakers have and give reasons for their judgements. Knowledge claims and beliefs often do stem from just such reason-based roots. This in turn motivates and explains behaviour; behaviour whose existence it seems equally pointless to deny. A speaker has good reason to claim that P when he recognizes the satisfaction of criteria for P and

the absence of any reason to doubt that P. Surely this is when many knowledge claims are actually made. Claims may even be made in the knowledge that the negation of what that claim asserts is perfectly conceivable. But if there is no present reason anywhere available to doubt the truth of the claim, is it in some sense a dubious claim because of that ? To return to the point, a criterialist will accept that S knows that P only if P is true, but will not accept that S never had any justification for the belief that P unless P is true. While it may be argued that a defective justification is no justification at all it must be noted that even defective justifications ground claims to know, motivate behaviour etc. The role of defective justification will, I think, be proved an important one, for even a defective justification may lead happily, though accidentally, to a true claim, as, indeed, a guess might.

It is in virtue of the fact that criteria are the learned paradigms of use and of their roles as reasons, justifiers and conditions of assertion that they do offer an account of why speakers claim, as they do, to knowledge, namely, on the basis of criteria which they learned when they acquired language; even though the correctness of their claims cannot be guaranteed. Secondly, nothing which has so far been said departs significantly from actual use or condemns in any sense speaker's understanding. In those cases where speakers clearly do recognize the satisfaction of criteria they will not hesitate to assert. They will unequivocally make claims to know, say, of pain. In so doing they will exemplify their linguistic training. Moreover, because criterial relations are conventional norms, to deny the conclusion of the criterially supported claim must be to possess an understanding of that

situation's aberration from the norm. Denial will require special justification. Here it begins to be clear why the theory has been thought to provide Sceptic-proof knowledge. The only doubt admissible as coherent in these situations will be grounded doubt and the burden of proof will be transferred to the Sceptic. Lacking altogether any such evidence, denial must imply either lack of understanding on the Sceptic's part or his mere objection to convention; to ordinary forms of language and discourse. In the total absence of contravening data, to refuse to assert is to reject linguistic training in the applicability of concepts.

While this clearly restricts or mitigates scepticism it does not amount to "Sceptic-proof knowledge"; but if we are to give up our ordinary ways of talking and reasoning, at least the onus is on the sceptic to give us a good reason why, on pain of being considered to have failed to understand linguistic goings-on. Here, in the ready transfer of the onus of proof, lies the anti-sceptical value of the theory. The fact that we do not have a final refutation of scepticism here, is not the fault of the theory but of those who issued extravagant promissory notes and claimed too much on its behalf.

Again, the claim is that if there actually is no concrete reason to doubt that P or to claim that -P then the sceptic is objecting to convention and rejecting his linguistic training. This is not to say that the negation of P is not imaginable, conceivable or logically possible. But is the very imaginability or logical possibility of -P itself a reason to doubt that P when P's criterial circumstances obtain ? If the psychological impossibility of my believing that P does not entail -P, then why should the psychological possibility of my imagining that -P, in the absence of ANY present reason, entail or even ground -P ? This point is reflected in the Dialogic [below Chapter 7] by the fact

that doubt, though an act, is not an assertion. The logical possibility of $\neg P$ is neither a proof nor a ground for $\neg P$. Just because $\neg P$ is imaginable in the absence of any reason to doubt does not mean that $\neg P$ is a reasonable basis on which to proceed. A society which did take this as a sound basis for action would be an irrational one far removed from our own, with concepts of knowledge and negation equally as distant. In fact there may well be problems of radical translation here, for that society's concept of negation is essentially different from our own. Rather than the familiar account of the negation of P in terms of it not being the case that P the alternative society maintains an account of P 's not being the case as equivalent to it being impossible that P . The modalities of such a society might be a perverse realization of a non-standard modal system, but it would, for all that, be perverse. If our hypotheses, beliefs and knowings are to be any aid to our survival then, I think, that society might be a dangerous and perhaps short-lived society. As I have implied, it is one which we cannot even fully make sense of. In general, I think it would be acknowledged that when listening to a recording, an actual scratch sounds significantly worse than several logically possible scratches.

At its strongest then the criterial claim is that a claim to know is, by definition, reasonable, just in those cases where the criteria obtain and is not reasonable where a genuine defeater obtains. This claim is as strong a claim as I wish to make at this juncture, urging that the first feature be replaced by the notion of criteria as learned norms, and by appeal to the distinction between the imaginability of the negation of the criterial claim and the existence of a concrete ground to doubt. If we characterise the simple objection as the following conditional: 'if

recognition of the satisfaction of criteria is always consistent with further information turning up which defeats that claim then satisfaction of criteria just cannot add up to knowledge' then, it might be argued, the conditional fudges an important matter and must be denied. That is, although it is fair to say that further information could turn up to defeat the assertion, this is to say no more than that speakers make criterial assertions against an explicitly understood background of defeasibility. It does not follow from this that it will always be rational to expect that one's assertion be defeated. In fact in many cases that would be quite irrational. My argument is that in the type of situation in which the burden of proof is transferred to the sceptic, the very last thing to be expected is that the criterial assertion be defeated.

A different response might however be made, and it is one which it is worth pausing to consider. The response is exemplified, for example, in P. Klein's [1981] "Certainty: A Refutation of Scepticism" which attempts to reconcile defeasibility and knowledge. Klein's account is not couched in criterial terms, but by giving it a criterial reading the plausibility of his approach might be tested in the light of the present considerations. In terms of the conditional I have been discussing, Klein's argument would proceed by agreeing with the antecedent: that recognition of the satisfaction of criteria is always consistent with further information turning up which defeats the claim. Klein would however deny the consequent: that recognition of the satisfaction of criteria cannot add up to knowledge. The crux of Klein's claim is that the absence of defeaters or overrides and the absence of any ground to doubt or the presence of only irrelevant or non-genuine defeaters, together with full criterial justification, ensures that P is certain on

the basis of the criterial evidence. In other words if we add to the satisfaction of the criteria, as a premise to the argument, the condition that in the present circumstances no defeaters actually obtain, that, it is held, will be sufficient to entail the conclusion of the inference. That is to say, if it is known that the criteria are satisfied and that defeaters are absent then, Klein argues, it is known that P is true. There is nothing to override it, nothing to defeat it and not even any ground to doubt it. This claim is bound to raise a number of objections not the least of which will be that despite earlier claims have criteria not now been accorded the status of entailments ? But Klein claims that this is not so and warns against such an interpretation in the following way. Even in cases where defeaters are absent it is not that nothing could defeat the criterially-based assertion. That would make the descriptions of the satisfaction of criteria themselves entailing. The point, Klein argues, is that nothing defeats the assertion in the present context; as a matter of fact. The absence of defeaters is the absence of ways to falsify. If the claim has not been defeated it has not been falsified. If there is nothing in the set of evidences which could defeat it, then it cannot on that basis alone be falsified.

This then is Klein's [1981] proposal. But it is difficult to see why this is a solution rather than itself a problem, for the guarantee of the absence of defeaters which was introduced as a premise in the argument is not, and could not be, itself another set of evidence, access to which could provide Cartesian certainty. What counts as a defeater could not be simply specified in a finite list in advance of a context. That nothing will turn up to defeat a claim in a given context is not something upon which one could base a claim. It is a wholly extrinsic factor. For the

moment I will hold my doubts in abeyance so as to further examine Klein's proposal.

By Klein's [1981] 'criterial model of knowledge' we know that P when: (1) We have a full non-defective justification of P and (2) There are no defeaters of P. But, if $\neg P$ were the case we could conjoin $\neg P$ to the left hand side of the inference to P and thereby defeat the justification, making it both defective and defeated. So, if P is non-defectively established and there are no defeaters of it then $\neg P$ is ruled out because it is itself a defeater which would generate a defective justification of P. Therefore, $P \rightarrow \neg\neg P$. But the evidence will still not be entailing because what rules out $\neg P$ is just the absence of any defeater, NOT the evidence itself. The assertion based on criterial evidence itself remains perfectly defeasible in practice, not just in principle. Any criterial assertion is always defeasible by a relevant set of defeaters. But where every member of that set is absent it is undefeated, P is established and $\neg P$ is ruled out. Now, even though $P \rightarrow \neg\neg P$, P itself is still not entailed. According to Klein that is just to fudge the distinction between defeating the inference to $\neg P$ and establishing that P. Hence, it is held: (1) To establish that P is to defeat the inference to $\neg P$; for it could then be conjoined to the left hand side of the inference. Of course, it is not being held that, (2) to defeat the inference to $\neg P$ is to establish that P. The fact that the inference to $\neg P$ is defeated might support X, Y, Z or nothing at all. It may bear no relevance to P. So, even if we conclude $\neg\neg P$ on the basis of evidence of defeat we cannot validly conclude that P. So, it is not a theorem that $\neg\neg P \rightarrow P$. That is to say, P is not entailed; its status

remains not proven. In fact, if we run the same initial argument for P, for -P, we get : (1) -P is fully justified. (2) There are no defeaters. So, (3) P is ruled out, ie. $\neg P \rightarrow \neg P$. This is the criterial consequence. Therefore, to know that P is to know that -P is not the case, ie. $P \rightarrow \neg\neg P$. To know that -P is to know that P is not the case ie. $\neg P \rightarrow \neg P$.

Presumably, it is not being held that if P is known then necessarily P is the case. P does not become a necessary truth in virtue of being known. Nor can it be held that if P is known then necessarily nothing could defeat P. That would mean that P was entailed whereas, criterially, there are always possible defeaters of P. But because P cannot be known ie. justified and undefeated AND defeated at the same time - $M (P \text{ is known} \ \& \ (\exists d) \ d \text{ defeats } P)$. Substituting -P which is a defeater of P for the second conjunct in the sentence above: $\neg M (P \text{ is known} \ \& \ \neg P \text{ is the case})$. Hence, $L (\text{If } P \text{ is known Then it is not the case that } \neg P)$ or $L (KP \rightarrow \neg\neg P)$.

The problem with all this is that the entire edifice of argument depends upon the notion of a 'clear cut case' in which not only are there, as a matter of fact, no defeaters but this is known to be so. While Klein's approach tries to preserve the defeasibility of the evidence it leaves the status of this extrinsic knowing of the absence of counterevidence very problematic. How do we know that this condition holds ? Klein admits that in the last analysis accepting that there are such cases, or accepting that THIS is such a case may not even be a matter for argument. Klein does argue that pointing to examples from everyday experience is persuasive and that it is more counterintuitive to deny that there are such clear cut cases. Despite this the worry clearly remains that there is nothing to prevent iterative scepticism of the form

: do we ever know that we know ? from running rampant here. I have already argued that we cannot just introduce the absence of defeaters into the argument as a premise and, now it seems that it is only if we accept an article of faith, accept that it is certain that no defeating evidence will turn up, that we can claim certainty for the conclusion of the criterial inference. It is only by assuming that we can know of the absence of counterevidence that we can argue that the criteria can give us knowledge. At the end of the day, the most basic foundation of Klein's argument is not something which is either justified or justifiable.

Klein is by no means alone in his attempt to reconcile the defeasible nature of criteria with the highly plausible claim that if S knows that P then P must be the case. Another attempt at such a reconciliation is contained in Gordon Baker's [1974] article on criteria in "Law, Morality and Society". Although there is no presently available formal characterisation of criteria, Baker's paper is very near to the mark here as it contains definitions of a degree of precision which easily give rise to what would appear to be the most obvious route to formalization. Acknowledging what has gone before Baker argues that criteria do provide sufficient evidence for assertions of pain, proposition P, unless some members of a set of defeaters of P obtain, let us say pretence or mimicry criteria. It appears that there are a number of possible symbolic representations of this claim that criteria, together with the absence of defeating evidence, warrant P. For instance,

1. $X, -D \vdash P$
2. $X \vdash (P \vee D)$

The first of these sentences might be taken to represent Klein's

[1981] proposal. However, I think that, strictly speaking, Klein's proposal is better represented as 3. $X, -D \vdash \neg\neg P$. The two sentences 1. and 2. are however logically equivalent if the entailment is taken classically, and I note that, in the latter case, even if X holds both P and D could be true in virtue of the inclusive nature of logical disjunction alone. Due to the equivalence, of course, if the latter will not do then neither will the former. This is itself an undesirable consequence, but the situation is worse still. In fact, X and D as they occur in these two sequents must, ultimately, be interpreted as sets. Here X is the criterial set for P , D a set of defeaters, and P our pain proposition. Both sets X and D are themselves infinite sets. The set X of criteria for P and the set D of defeaters of P will be incompletely specified, just because we cannot know in advance exactly what will count as a criterion or what will count as a defeater in any situation. These will not therefore be closed lists. Criterial inference cannot therefore be formally characterised in that way as neither X nor D could be well-defined. Therefore any adequate formalization must make appeal to a semantical, model-theoretic notion, for only in that framework can infinite sets be unproblematically accommodated.

Baker is aware of the problem presented by these infinite sets and attempts to overcome them, for D at least, by arguing that the defeaters will at least be of distinct types and this will provide a finite list. Hence, formally:

Let A_1 be a set of defeaters of a type: $\{a_{11}, a_{12}, a_{13}, \dots, a_{1m}\}$

Let A_2 be a set of defeaters of a type: $\{a_{21}, a_{22}, a_{23}, \dots, a_{2n}\}$

Let A_k be a set of defeaters of a type: $\{a_{k1}, a_{k2}, a_{k3}, \dots, a_{kp}\}$

I take A_k to represent some finite cut-off point in the list.

Hence,

we have the set: $A_1 \cup A_2 \cup A_3 \dots \cup A_k$

Now, something is a member of the original set of defeaters "D" in our sequents only if it is a member of the union of sets of types of defeater. ie. $\varphi \in D$ iff. $\varphi \in A_1 \cup A_2 \cup A_3 \dots \cup A_k$ ie.

$$\varphi \in D \text{ iff. } \varphi \in \bigcup_{i=1}^k \{A_i\} = A$$

With the two sentences reformulated set-theoretically, Klein's sentence can now be read as 'C' & $\neg(\exists q) q \in A \vdash 'P'$, where the turnstile represents semantic entailment. The idea then is to treat the new sentence as meaning that there is a finite subset giving you an inference ticket to P. That is, P is deducible from the finite subset. But entailment is not well-defined here and really ought to be semantic. However, once it is realized that in fact A must itself be an infinite set if any sense is to be made of it, then it can be seen that even the finite set of infinite sets is of no real help, because it must be the case that all of its members must fail to hold. Further, because we cannot know in advance what individual members of any particular set might be, each individual set will itself be incompletely specified ie. the sets: $\{A_1\}$, $\{A_2\}$ to $\{A_k\}$ remain incompletely specified. This cannot provide a solution. In the last analysis to accept the infinite set at all, to speak of it as not holding in a given case and so on, is all on a par with the notorious notion of a ceteris paribus clause for law-like statements in scientific contexts, and that is surely just the type of move which Klein invited us to make by taking on board, as an article of faith, that such and such a case was in fact a clear cut case. I conclude

therefore that neither Klein [1981] nor Baker [1974] takes us any further forward towards an answer to Wright's [1984] simple objection than the semantic theory of non-monotonic inference did.

Interestingly, under the Klein/Baker scheme just discussed the same type of construction must be made for the set X of criteria for P and it is very far from obvious that such cumbersome and incompletely specified definitions could ever satisfy the condition of teachability. If this were the form of any initial definition or explanation of meaning would that be a teachable form? The sets are unspecified and unspecifiable for every situation on the basis of any single case. Though that qualm might be overcome in the last analysis it does not augur well for proponents of any such view, and I do not believe that this type of account accurately represents how criteria are learned. They must be learned as more simplistic inferences understood against a background of defeasibility. Precisely what counts as defeating evidence must also be learned. The complexity of possible defeaters may be less immediately grasped than the simple rule and degree of appreciation of social, and evidential cues to asserting defeat may develop as language-mastery deepens. But that complexity can never be completely specified in advance in any learnable definition. It is not obvious either that this type of formal framework has enough scope to do full justice to the finer points of criterial reasoning. Finally, there remains a more fundamental reason why none of the attempts to reconcile criteria as defeasible evidences with objective knowledge could ever be made good and that reason will become clear in the following section.

III DEFEASIBILITY AND THE MEANING FEATURE

So far in the present chapter I have discussed issues arising from

the first challenge to the coherence of the concept of a criterion as described in Wright [1984]. In the light of that discussion I have relinquished the first, "knowledge feature", on the grounds of its indefensibility. In addition I have argued against the claim that such a feature has a genuinely Wittgensteinian pedigree and have maintained, against current orthodoxy, that the acquisition of sceptic-proof knowledge is not the main point of the concept of a criterion. In place of the knowledge feature I have argued for the recognition of the satisfaction of criteria as a reasonable basis for knowledge claims and have urged a distinction between the imaginability of the negation of P and the existence of a concrete ground to doubt P, which, I hold, provides the basis for a dialogical transfer of the burden of proof to the sceptic. In the sections below I shall try to give some further substance to these views while addressing the second major objection to the coherence of the concept of a criterion presented in Wright [1984]. That objection concerns an alleged incongruence between the defeasible characteristics of criteria and their role as content-fixers.

My strategy will be to address each point to which I will object, as these arise in the course of the article. Over pages 291-293 Professor Wright explains that because we cannot test anyone's understanding of statements whose truth-conditions lie outside our recognitional grasp by reference to those very truth-conditions, the anti-realist is faced with the dilemma of either giving truth-conditions which we can so grasp or admitting that:

"...for these statements he is well advised to prefer the second response: their meaning is explained by reference to conditions of warranted assertion whose obtaining is not sufficient for their truth."
(p.393)

Further, it is held,:

"..for the assertion-conditions theorist every contingent statement must be associated a priori with at least some of its assertion-conditions.." (p.393)

It will be clear from my Chapter 4 that the enterprise of constructing a global theory of meaning conceived in realist or anti-realist terms is not the purpose of my own defence of the coherence of the concept of a criterion. Nor do I wish to lay claim to having anything to say about "every contingent statement". Therefore, before going any further it may be prudent to make clear my own position, particularly as such notions as 'testing for understanding' and 'explaining meaning' are under discussion. My own position here is a logical development of some of Wittgenstein's arguments in the Philosophical Investigations. As a logical development however I take full responsibility for them and would be extremely reluctant to attribute my own account to any other author. My own view is a logical development in so far as it proceeds from Wittgenstein's own distinction, familiar from Chapter Two, between language which pertains primarily to public physical objects and that pertaining to sensation. In many of his remarks in the Investigations Wittgenstein is at pains to emphasise the distinction between the bearer of a name and the meaning of a name (ie. 40-45), or between what a word means and what it refers to. Despite this, that author does acknowledge that the referent, the bearer of the name, can be used to assist in teaching the meaning of the word. In Chapter Two I argued that Wittgenstein was at pains to emphasise the important differences between the grammar of the public physical object language-game and that of psychological language. In particular, I discussed this issue with reference to Wittgenstein's claim that a sensation was no sort of object but was only considered as such when the

grammar of the language of sensation was confused with that of physical object language. Despite his denial of the Cartesian, and indeed the Classical Empiricist, conception of sensation I argued that this did not imply that Wittgenstein was espousing the view that what the language of sensation referred to was, purely, behaviour. I denied, with Wittgenstein, that a sensation, if not a something, must therefore be a nothing. Rather, Wittgenstein tried to cut a middle path between behaviourism and the object conception urging that both were mistaken. In the present context what the object theorist has got right is that the language of sensation refers not to behaviour but to sensation and that it is the presence or absence of the sensation which settles questions of truth and falsity. Where a mistake is made is in thinking that the psychological phenomena referred to could be used to teach or test the meanings of those referring terms. The behaviourist makes precisely the opposite mistake when, accepting that behaviour has a vital role in teaching and testing the use of the language of sensation, he moves to argue that sensations are nothing other than the behaviour. The way in which a middle path is cut out here is to accept from the object theorist that the language of sensation refers to sensations understood as psychological phenomena, while accepting from the behaviourist that such phenomena cannot be used either to teach or test meaning. If the meanings of sensation terms cannot be taught by reference to the psychological phenomena then they are taught by reference to something more like behaviour. I say 'more like behaviour' because I do not wish to confine myself only to behaviour, be it linguistic or not, but want to include circumstantial and contextual features and, in extending the analysis to emotion, to allow a role for the object of the emotion, should one exist.

In other words, it is by means of criteria that the meanings of psychological terms are taught despite the fact that it is the psychological phenomena, and not the criteria, which our use of psychological language is about.

If we now return to the original distinction between psychological terms and public physical-object language the vital difference between these two ways of talking, I argue, lies at the level of basic teaching and learning. That is, while in the case of the language of physical objects the object itself might enjoy a useful role in the teaching of the meaning of the word to which it stands as referent, this is precisely what is not possible in the case of psychological language; not least because the referent in these contexts is never any sort of object at all. There is, therefore, a fundamental reversal of logical priorities here; for while in the public object case the referent can be used to establish, illustrate and teach the meaning; in the psychological case the meaning, given by the criteria, is used to establish the referent, even though the criteria are not that which is referred to. However, none of this implies that our talk about toothache is about toothache-criteria rather than toothache. The meaning of 'toothache' is given by its particular criteria and those criteria fix what toothache is. It is in terms of criteria that we fix the referent of that term, and in terms of criteria generally that we identify, distinguish and classify the meanings of the psychological terms we use. If toothache were not in general expressed as it is then, whatever the feeling, it would not be toothache. Furthermore, our talk about toothache cannot be reduced to talk about anything else. What makes a particular assertion of toothache true is not just facts about toothache-criteria as distinct from facts about toothache. For the facts about the criteria for toothache are

themselves facts about toothache. Our criteria for toothache are criteria for toothache relative to our concept of toothache. Even in the first person we acquire our linguistic understanding of what our feelings are and classify and discriminate them by means of the criteria which fix the meanings of the terms we learn. Later, of course, we need not identify a particular sensation by, say, observing our behaviour, that would clearly be absurd; but that does not imply that our understanding of the meanings of the psychological terms we use, even of ourselves, is not, ultimately, given to us in criterial terms. This is the whole, holistic, story and there can clearly be no other mode of access to our inner life, in so far as that life is conceptualised, other than in terms of an understanding of the language of the inner life which is taught and learned on the basis of criteria. Our talk about toothache criteria is, therefore, in a deep sense, talk about toothache. Further, talk about the satisfaction of toothache-criteria is not ultimately distinguishable in any sense from talk about toothache for it is the criteria which determine the meaning of the term 'toothache' and fix what toothache is. This rather holistic, non-reductionist position is by no means the only possible position but is, I think, seen to be the most plausible when compared to alternative accounts.

That is, if we separate the criteria and the phenomenon and argue that facts pertaining to the phenomena settle the issue of truth or falsity rather than facts about criteria, if we argue that, in some sense, these are distinct, then it looks as if statements might be true in virtue of the obtaining of a condition that outruns the criterion which determines the very meaning of the term. Again, use in accordance with meaning and saying something true are presumably inseparable.

Otherwise, it looks as if there might be a further question as to whether a criterially warranted assertion is in fact a true statement. What is now needed in this particular case is an account which allows for a conceptual difference between criteria obtaining and truth-conditions, strictly speaking, obtaining; but an account which yet explains how the notion of truth here is nonetheless constrained by that of criterial assertability. Any criterialist who wanted to pursue that particular line must make fully explicit the precise way in which:

"...the truth of a statement may be connected importantly with the truth of another without it being the case that the one entails the other in the sole sort of sense preferred by obsessional logicians."
[J.L. Austin, 1955, "How To Do Things With Words", p. 54. Austin's remark, of course, concerns presupposition and not criteria.]

Part of any such successful account must involve driving a wedge between 'use in accordance with meaning' and 'making true statements'; at least as far as criterially-based language goes. But there is clearly a grave problem with that very idea, for how can the criteria determine the meaning without fixing the referent of the term in question? There is clearly a grave danger of cutting adrift the psychological phenomena to some ghostly realm and attendant upon that scepticism about other minds. It could be argued, for instance, that it is perfectly possible both to use language correctly and meaningfully and yet to say something false in the process. But that fact is surely contingent upon the meaning of the term establishing its referent in such a way that we already have a basis for knowing when the term is used truly and therefore when it is used falsely. Rather, I argue, in the case of criterially-based language, the criteria give the meaning of the term and it is only via that established meaning that the very nature of the referent is established. Of course, that conclusion must cast doubt on the view that

it is really possible to give any account of the content of criterially-based assertions about mental states from outside that very type of language. Further, because the problems with a non-monotonic account of truth remain unsolved the core notion in the semantics of psychological language must be illustrated and explained by the interaction of criteria and defeaters. This at least allows us to map reasoning patterns via something more like assertability and non-assertability rather than truth.

It is important to note that the view which I am advocating here does not necessarily imply that we had no pre-linguistic psychological phenomena. These will have existed as, unconceptualized, "blind intuitions" in Kant's terms. The criteria give us ways of carving up, discriminating, classifying and conceptualizing our experience. Similarly, it is vital to note that criteria cannot be considered evidence in the ordinary sense. For if criteria were evidence in the ordinary sense then we ought to be able to provide a characterisation of the referent which is ultimately independent of its criteria. But that is precisely what is not possible. No independent characterisation is possible, for it is only by means of the criteria for X that we can understand what X is. This point in turn undermines the degree of depth claimed for the first person / third person asymmetry in these contexts. To understand what an X is for the psychological case presupposes understanding of the criteria for X and this holds regardless of perspective. Of course, once this much is understood one can identify X in one's own case without explicit appeal to criteria, but it is only in virtue of the criteria that one can conceptualize from the feeling to the feeling as an X. It will also be noted that on my view avowals, sincere

or not, cannot be construed as criteria. This, of course, stands in stark contrast to the view usually attributed to Wittgenstein that such avowals could form part of a criterion for the ascription of mental states to others. I do not however wish to attribute my view to that author. Correct use presupposes knowledge of the criteria in terms of which the correct use was taught, while incorrect use simply exemplifies lack of understanding. The desire to treat the product of having learned criteria, correct use, as itself a criterion is an effect of the desire to see criteria, fundamentally, as evidence in the usual sense.

These conclusions, I argue, must give the lie to the type of view which Paul Churchland [1986] has recently put forward in the *Philosophy of Mind*. I have already critically examined the presuppositions of that author [above, Chapter 4] now I will consider his conclusions. As is well known Churchland heralds a great enhancement in our 'introspective vision':

"Glucose consumption in the forebrain, dopamine levels in the thalamus, the spiking frequencies in specific neural pathways, resonances in the nth layer of the peristriatal cortex, and countless other neurophysiological and neurofunctional niceties could be moved into the objective focus of our introspective discrimination and conceptual recognition..We shall of course have to learn the conceptual framework of the projected neuroscience in order to pull this off. And we shall have to practice to gain skill in applying those concepts in our noninferential judgements. But that seems a small price to pay, given the projected return."
(1984, p.160)

But how much of an advancement would our using the vocabulary of a completed neuroscience about emotions and mental states really represent ? I would argue that Churchland is simply advocating a change in the way we talk based on our original emotion or mental state vocabulary whose meanings are criterially given. If the taxonomy of folk psychology is to be replaced, it must, as Churchland realizes, map onto the taxonomy of

the completed neuroscience. But that is precisely because our ordinary ways of talking and our common-sense explanations fix the meanings of the terms, defining what each state is, and therefore the science must be associated with that if we are to understand the new science-talk as talk about emotions. In that sense, the new way of talking is parasitic upon the old, inescapably tied to the carving up, the discriminating and classifying already done. While it might represent a new way of talking, it can be no more than a new way of talking about the same old thing. It is very difficult to see how a change based inescapably on what has gone before and only comprehensible on that basis can really amount to the great revelation Churchland foresees.

With this brief statement of my view, in contrast to the context in which Wright [1984] embeds the concept, I return to the issues raised in Wright's article. One passage which brings into relief the contrast between what I take to be the correct conception of a criterion and the conception discussed by Professor Wright is found on p.390:

"The crucial consideration is again the consequential character of assertion..when someone asserts P, even on inconclusive grounds, he sets himself against the subsequent defeat of those grounds. Nevertheless, defeat always is a possibility where criteria are concerned. And it will be in the lap of the gods whether it occurs in any particular case and how often it happens..It is to be expected no doubt - if only for evolutionary reasons- that we will in general have selected criteria so as to minimize the possibility of defeat. But that is not to say that we have been successful in any particular case, or that success will last. So..it has to be a possibility that a type of criterion for a particular assertion be defeated often enough to shake our confidence in the propriety of that assertion when made solely on that ground."

The same point is stressed on p.398:

"For anyone who understands both that a certain type of state of affairs conventionally supports P and that this support is defeasible by developments which, again, he understands..must understand how experience could lead one to the conclusion that occurrences of that type of state of affairs did not after

all provide an adequate reason for believing P. The possibility involved is, once again, that of the world turning awkward... worldly awkwardness may precisely take the form of an erosion of those conditions."

The general worry here stems, firstly, from the fact that criteria are still construed as primarily evidential and, secondly, from too broad, too global, a conception of criteria. If we narrow the focus to psychological terms, I argue, the anxiety about the erosion of criterial connexions looks much less well-founded.

What has to be borne in mind, I argue, is that unless the criteria for pain were as they are then there would not be pain. If there were no such criteria there would be no pain. That is not to deny the pre-linguistic existence of psychological phenomena but to point out that these phenomena would, ex hypothesi, remain unconceptualized, again in Kant's terms the intuitions would be hopelessly blind. Further, if there is no independent means of characterising the psychological phenomena other than via the criteria in virtue of which the terms referring to the phenomena can be understood at all, then if the criteria for pain goes, so does the pain. And that is exactly what we would expect from the type of relation which a criterion is supposed to provide. Therefore, in general, recognition of the satisfaction of criteria for pain must be the basis of correct use of the term pain. That is why, although any criterially-based knowledge claim about pain might be defeasible, that is no cause for anxiety that one particularly awkward day the criterial circumstances of pain might no longer have any relevant connection with pain. Without the criteria for pain there would, literally, be no pain. Again, it is not that the agent has no reason to doubt the criterial connexion in general but that he could have no reason to reject the criterion per se unless his purpose was simply to alter those conventions

which determine the meaning of the relevant term.

In "Criteria, Defeasibility and Knowledge" [1982], mentioned briefly above, John McDowell outlines an alternative approach to the whole concept of a criterion. In that paper McDowell opposes the orthodox interpretation of Wittgenstein's later work as featuring a notion of criteria as intrinsically defeasible. Further, McDowell sketches a solution to the epistemological problem which I have discussed in the present chapter. In the first instance McDowell [1982] replaces defeasibility with a disjunctive thesis in the philosophy of perception to the effect that, either one perceives a mere appearance, or experience discloses a fact in perception. If the judgement on the basis of a (mere) appearance falls short of the fact then that judgement simply was not a genuinely criterial judgement. In allowing as genuinely criterial only judgements about acts of perception in which experience actually discloses a fact McDowell ensures that his criteria fix reference and that correct use on the basis of criteria is identical with the making of true statements. However, as indicated earlier, the price to be paid is the loss of defeasibility. McDowell, in effect, replaces defeasibility with the claim that ascriptions of mental states are decidable perceptual statements and therefore bivalent. It is doubtful however whether this really solves any of the deeper problems raised by criteria, and indeed doubtful whether McDowell's view can really stand on its own legs at all. That is, in making mental state ascriptions simply decidable, it could be argued that McDowell, far from addressing the problem, merely circumvents the problem by shifting the borderlines between the decidable and the undecidable. For Dummett, for example, nothing is gained as regards the problem of the notion of truth relevant to undecidable propositions. McDowell has avoided, rather than

answered, the question. The disjunctive thesis, it is claimed, is however of genuinely Wittgensteinian pedigree unlike that of defeasibility. McDowell motivates the latter claim by appeal to remarks 377 and 354/5 of Wittgenstein's "Philosophical Investigations". The upshot of remark 354, McDowell argues, is not that the criteria for rain are satisfied even though it is in fact not raining but rather that the criteria never were satisfied, they only appeared to be so. The same reasoning is applied to the idea that the possibility of pretence implies that criteria must be defeasible:

"Here is a possible alternative; in pretending one causes it to appear that criteria for something 'internal' are satisfied..but the criteria are not really satisfied."
(p.466)

Remark 354 is an unfortunate choice upon which to ground an interpretation of Wittgenstein for it is far from uncontroversial. In his "Wittgenstein's Private Language Arguments" [1986], for example, Bernard Gert argues that the remark is both confused and mistaken in suggesting that private sensations could function as public criteria. But I do not accept that this is the point of 354. Rather, remark 354 suggests the more interesting thesis that sense- impressions are contained inside language and do not provide a neutral ground or justification for language. Remark 377 does however provide better support for McDowell's alternative. This remark, I believe, points up another consequence of the distinction between psychological language and physical-object language, namely that in language-acquisition objective language is logically prior to the language of experience. In remark 377 Wittgenstein contends that the distinction between "being red" and "seeming red" can only be taught at all if the concept "red" has already been understood. The objective is the 'first' language-game. Notably,

this reverses the logical priorities of Cartesian epistemology for what is indubitable is not, by that fact alone, foundational. The language of experience emerges on this account as parasitic upon an objective language-game which has logical primacy. This is a better ground for McDowell's alternative reading of "criteria" though it cannot be a conclusive one. Despite that fact the onus is on the supporter of defeasible criteria to produce textual evidence of Wittgenstein's commitment to defeasibility; a difficult task given the scarcity of such evidence.

Finally, McDowell [1982] is at pains to cast doubt upon the view which compares criteria to "highest common factors" ie. features which are present both in actual cases of criteria and in cases which only seem criterial. It is in virtue of just these shared features that we can take an instance of the latter as being of the former. Now, although we are warned against this view it may have more plausibility than McDowell allows. Further, an examination of the plausibility of that view may cast doubt upon the viability of McDowell's position.

In the first section of his paper McDowell [1982] outlines an alternative to the criterial position; an alternative which is labelled "M-realism" by Crispin Wright [1980]. In order to articulate M-realism McDowell quotes from his earlier paper "On 'The Reality Of The Past'" [1978] as follows:

"...we should not jib at, or interpret away, the commonsense thought that, on those occasions which are paradigmatically suitable for training in the assertoric use of the relevant part of a language, one can literally perceive, in another person's facial expression or his behaviour, that he is [for instance] in pain, and not just infer that he is in pain from what one perceives."
(p.456)

If we take this basic account of M-realism together with the

disjunctive thesis then we must also accept that 'paradigmatic suitability for training' is determined by it being a fact that another person is in pain. Otherwise, by the disjunctive thesis, the judgement would be based on mere appearance, could not be criterial in McDowell's [1982] sense and therefore could not be paradigmatically suitable for training. But this implication conflicts with the commonsense idea that one can illustrate and explain the criteria for a psychological term in the absence of the psychological phenomenon itself. That is, it conflicts with the intuition that it is at least unnecessary and probably counter-productive to become sad in order to teach the meaning of sadness, or to be in pain to teach the meaning of the word "pain". But once that is conceded the criteria as determinants of meaning must be considered apart from the referent of the term whose meaning is so determined. In other words the 'highest common factor' conception (and indeed defeasibility) come back into the picture unless we deny the very commonsense intuition about training in the use of psychological vocabulary.

IV CONVENTION AND CONTINGENCY

Prof. Wright's [1984] paper closes with an appeal for clarity and precision in the form of a challenge: if ($D_1 \dots D_n$) criterially support P when are the substitutions on D necessary and when contingent ?

Any serious theory must attempt to meet this challenge though, I think, the terminology will not prove useful here. If $D_1 \dots D_n$ is to represent criteria as those situational and behavioural circumstances in virtue of which we learn when to assert P , when to say that P , and determine what we mean by P then the schema is best understood as

representing a rule. In so far as that is the case the schema represents a grammatical rather than a simply empirical fact. If this is a necessary connexion it is necessary in that the connexion is semantic and the relation content-fixing with respect to P, but in a sense which is not expressible in terms of formal logical notions. In Wittgenstein's terms, it is a necessity of our form of life that certain kinds of behaviour in certain contexts are criterially related to pain in that it reflects a basic fact of the human condition. This is, perhaps, not even to say that it could not have been otherwise; but the fact that it is so represents a universally-shared feature of human life. While alternatives might be imaginable or conceivable this does not imply that we could ever actually live by them. We plainly could not legislate away pain-behaviour and enforce a new way of acting. This fact, I argue, must cast some doubt upon the idea that the putative alternatives are completely and coherently conceivable. If we take seriously for a moment the sceptical doubts and worldly awkwardnesses outlined in Wright's [1984] second objection and try to conceive the resultant world in which none of our psychological phenomena were defined in terms of their ordinary criteria, where the recognition of the satisfaction of pain-criteria no longer provided a basis for the correct use of the term "pain", we will find it a remarkably difficult feat of conception. Is it, for instance, to remain the case that psychological terms be teachable only via criteria or not? Would it be that our usual psychological phenomena were replaced by phenomena of an unfamiliar sort? This is hardly traditional logical necessity but it does cast doubt upon the very conceivability of the kind of worldly awkwardness mentioned.

Any world with different connexions of this basic sort would be a

very different world from ours indeed. It is not any more useful to consider the schema as a convention either; for it is only a convention in the sense that a whole language is conventional. While grammar could be described as a convention many things which are called conventions have no bearing at all upon grammar. There is no useful analogy here to a congress of Physicists deciding to adopt an international symbol for force or mass. Facts of grammar do not always admit of this kind of free choice and blatant arbitrariness. The whole game of the avowal, assertion and attribution of pain and the use of the criteria for that term to explain, justify, learn and teach the appropriate performance of the acts in question is more like an inheritance; a taught linguistic heritage to which human beings have a basic susceptibility and markedly similar responses. The relation between the criteria for the term and the use of the term is not an inductive relation but an internal, semantic and content-fixing relation. In short, our particular criterial rules are contingent to the extent that we could have used language in another way; but they are necessary given the way we do speak. As stated above, it is not obvious that we could after all alter our language-game significantly. At least, I argue, we cannot fully conceive of what that would be like.

In certain contexts, I argue, facts of the human condition limit and determine the degree of change which is possible. Obviously, neurophysiological facts about human beings, pain and pain-reactions have a role as regards the language-game with the term "pain", and it is facts of this type which in a sense precede our form of language (or which are structure-independant, if I may misuse a Chomskian term), which criterial relations exploit and reflect at the linguistic level. The very fundamental character of these facts about our nature does clearly point

up a constraint which, I argue, must mitigate the sceptical concern that the changing world might ultimately undermine the nexus of criterial connexions in important ways. Any degree of change sufficient to wholly undermine the criteria for pain actually being a basis for the correct use of the term "pain", if that possibility is a coherent one at all, would clearly require very fundamental alterations in the physiological fabric of human life and would be tantamount to a serious change in the nature of the human condition. The existence of a working logic of criteria in everyday explanation and the success of criterial inference testifies to the unlikelihood of change to the hyperbolic extent imagined by the sceptic. If criteria are unreliable how is learning and teaching via examples even possible ? Again, these facts testify to stable connections between language and the world.

The schema also represents a pattern of explanation. The sceptic is correct in a sense when he asserts that it is possible that $\neg P$. But to fail to see that when $D_1 \dots D_n$ obtain, "P" is appropriate is to fail to see that $D_1 \dots D_n$ fix what "P" means. It is just to fail to understand P. At the meta-level the sceptic argues that it is always possible that $\neg P$. Given grounds the sceptic can doubt any particular assertion, but nothing could constitute a ground for doubting the mode of explanation. What could count as a falsification of recognition of the satisfaction of pain-criteria being the basis of correct use of the term "pain" ? Could everyone always be pretending ? To pretend to be in pain one must first learn what it is to be in pain, that is the semantic priority. There also exist criteria for pretending and while pretence certainly defeats the assertion of pain it never undermines the criterial connexion, rather it exploits it, or gets off on the back of the criterial connexion.

PRETENDING to be in pain is nonetheless pretending to be in PAIN. To refuse to accept criteria as explanations of correct use is to reject a standard model of explanation in the language we use. What is the alternative model of explanation which the sceptic has ? At least the onus is again on the sceptic, for by rejecting the rules he is refusing to play the game.

Although to some extent we may inhabit a world of Heraclitean flux we can talk about rivers and know what feet are. If the world changed or was in any danger of changing so much that criterial connexions could no longer be successfully maintained, if everyone in the dentist's was always and only missing games, then I think we might well not know what to say regardless of the theory we held. I have tried to show that what the sceptic invites us to imagine and to worry about, the coming to be of alternatives, may not even be coherently conceivable. The natural ebb and flow in language is I think in no danger of drowning us. Finally, perhaps the best way to characterise the schema is as a reminder of the rules of the game, a perspicuous representation, or in the dialogical sense a point of order. This is where justification ends which, I argue, is an appropriately Wittgensteinean response to a paper which in many ways champions the position of the sceptic.

CHAPTER SEVEN

INTRODUCTION

In the following chapter I attempt to represent the salient features of criterial reasoning formally. The formal system presented is intended as a clarification of the account developed in preceding chapters. The possibility of formalization will underpin the view that the notion of a criterion is not an incoherent one. Further, if the notion of a criterion can be made sense of, then, to some extent, the traditional views which oppose it will be undercut. Consistency itself is not proved here, although I hope to have transferred the burden of proof to the sceptic. The strategy itself is therefore one which coheres with the conception of criteria outlined in earlier chapters; especially chapter six.

I THE DIALOGUE SYSTEM

Criterial reasoning is represented within a context in a formal dialogue system. The system is designed to be a formal analogue of actual language-use so as to emphasise the salient features of the logic of criteria against a realistic background, which will permit the fullest examination of the use speakers make of criteria. A number of speakers are featured, two in the simplest case. The speakers are embedded in a context composed of behavioural and situational circumstances, which they can report to one another and exploit as evidence and as premises in arguments. In the type of theory to be considered criteria are taken to be those situational and behavioural circumstances of a linguistic act which make appropriate the performance of that act.

The motivation for the claim that descriptions of the satisfaction of criteria for P do not entail that P resides in the defeasible

character of assertions made on that basis. Criterial assertions must be defeasible if it is to be possible to account for acting, pretence, deceit and the like. Where Q, R and S are criteria relative to P it will always be possible that while Q, R and S hold, P does not. The criterial relation must therefore be a non-deductive relation distinct from entailment. Clearly, this defining characteristic of the criterial relation must be represented in any adequate formalization. In the present case an attempt to represent that feature is made by incorporating defeaters into a context and making them available to speakers. Here it is vital to note one advantage which flows from the dialogical framework and the notion of a context, which was not available on the Baker/Klein model: while it may not be possible to say in advance of a given context which defeaters might hold, that does not imply that in any given context the list of defeaters is always infinite, or that the question of whether or not a defeater holds is not decidable. I have argued above that the notion of 'defeasible knowledge' is one which should be replaced by that of a mechanism effecting the transfer of the burden of proof from the criterialist to the sceptic as the sceptic challenges a criterially-warranted assertion. The onus on the sceptic is to demonstrate in which respect this situation differs from the norm, that is, from those "paradigmatically suitable" cases on the basis of which his understanding of the assertion was given. I have also urged that the distinction between the imaginability of the negation of P and the existence of a concrete ground to doubt that P must be accepted on pain of simply rejecting the concept of negation. These two features at least must be captured in the dialogic.

The purpose of the dialogue is to correctly subsume the evidence

under a concept. By using the evidence and the criterial relations speakers conjecture, reason, falsify and seek to correctly interpret their environment. The range of linguistic acts available to speakers is limited to five: assertion, doubt, challenge, withdrawal and exhibition of evidence. Each speaker has a commitment store, initially a tabula rasa. The following sets are accessible to speakers:

1. A completely specified set φ of sentences. In purely criterial contexts these will all be criterially-based, ie. "X is in pain", "X is pretending". Asserted sentences of this form are, of course, understood to be defeasible, dependent upon the assignment of truth-values to the members of 2 below. That is, it is intended that the truth-values of these sentences be computable on the basis of assignments of truth-values in 2 below.

2. A completely specified set ψ of sentences intuitively understood to be criteria and defeaters. That is, these sentences make explicit the defeaters for the relevant criterial assertions in 1.

In general, the union of φ and ψ defines the context in which the speakers are taken to be embedded. The formal names of key terms and acts are as follows:

Speaker = anything which asserts, challenges, doubts etc. in a dialogue with appropriate commitment effects.

Context = the union of the contents of φ and ψ .

A -assertion operator, any formula constructed from statement letters together with the usual connectives.

D -doubt operator.

W -withdrawal operator.

? -challenge operator.

E -exhibition operator; technically, this is a move rather than an

act; a point to which I return below.

II RULES AND ACTS

The dialogue is composed of a language L , understood as the union of the contents of φ and ψ , and a set of rules. An atomic sentence in L is a sentence letter from the appropriate set of sentences. Complex sentences are formed from atomic sentences by combination with connectives.

For any point in a dialogue I give the general description:

(n) i : Op. (Sentence)

ie. Number of move (consecutive). Speaker (alternating). Operator.
Sentence.

Commitment is a function from the natural numbers and speakers to the power set of sentences: $C : N \times (A, B) \rightarrow P(L)$. Individual commitment at the n th move is a subset of the set of wff's of L , ie. a set of sentences. $C(n, i)$ is a subset of the set of wff's of L .

The rules for speech acts are as follows:

Assertion."A".

1 $(\forall n)(\forall i)$ If (n) i A(S) then $S \in \{L\}$.

This particular rule, that speakers can only assert sentences in the language, is not peculiar to assertion but holds for every operator. For that reason I omit to repeat it in each case, assuming that it is implicit in the characterisation of each operator. This allows concentration on the commitment effects pertaining to the use of each operator.

2 If (n) i : A(S) then $(\forall i) C(n, i) = C(n-1, i) \cup \{S\}$.

Acts are performed not asserted by speakers in a dialogue. Assertion

applies to sentences. In order to ensure that the dialogue will get off the ground I take assertion to commit all asserted statements to the commitment-stores of each speaker. These commitments can be removed by the speaker who has not himself asserted it, as I will make clear, but, in the first instance we assume everyone to be committed, pending a claim to the contrary. I do not believe that this is the only way to construe the commitment effects of assertion but it does force speakers to act against that commitment if they do not accept it and therefore facilitates an active dialogue and, hopefully, minimises the possibility of dialogues in which speakers assert sentences which just have no bearing on the commitments or acts of their opponents. Rule 2 ensures a certain common focus of argument and forces response on pain of the asserter's sentence being accepted by both, in the sense that both will actually be committed to that sentence. Any well-formed sentence may be asserted. Any act may follow an assertion.

Implied Commitment: if an assertion is made and followed by a challenge that must be followed in turn by an assertion. The second assertion is then understood to have been given as a reason for the first ie.

- (1) A : A(X)
- (2) B : ?(X)
- (3) A : A(Y)

Three step arguments of this form will result in the following pattern of commitment effects:

$C(1,A) = \{X\},$	$C(1,B) = \{X\}$
$C(2,A) = \{X\},$	$C(2,B) = \{X\}$
$C(3,A) = \{X, Y, Y \rightarrow X\},$	$C(3,B) = \{X, Y, Y \rightarrow X\}$

Intuitively, B has asked at (2) why A thinks that X is the case. A replies that Y, giving Y as his reason for thinking that X. Other rules

allow B to attempt to demonstrate the poverty of Y as a reason for X if he does not wish to accept A's argument. As most of the work done by criteria is in terms of their capacity to warrant an assertion, to justify or to provide a reason to assert, it is vital to give speakers mechanisms to make explicit the grounds of their opponent's assertions. Of course, we do guess, conjecture and make unwarranted assertions in ordinary discourse. It would be absurd to deny that. It would, however, be equally absurd to assert that we never have reasons for assertion. As the dialogic is to represent the sub class of grounds for assertion which are criterial in character it seems reasonable to endorse some such mechanism. Though there might well be a dialogic of Popperian conjecture that is not the primary purpose of my own dialogic. Further to the formal argument in three steps above however it is important to note that just 'if Y then X' will not do because A is not just saying 'if Y then X' but 'X is the case because Y is actually the case'. This feature of the formal dialogue points up the reason-based nature of assertion which, I have argued above, is integral to the logic of criterion. In the case outlined ' $Y \rightarrow X$ ' is an implied commitment. So, Ass. A(T) after B?(S) adds $T \rightarrow S$ to the commitment-stores of each speaker.

Implied commitments will also be generated when an assertion is given in response to a challenge as a reason for doubting a prior assertion ie.

- (1) A : A(X)
- (2) B : D(X)
- (3) A : ?(D(X))
- (4) B : A(Y)

The pattern of commitment effects generated by the four step argument will be as follows:

$C(1,A) = \{X\},$
 $C(2,A) = \{X\},$
 $C(3,A) = \{X\},$
 $C(4,A) = \{X, Y, Y \rightarrow -X\},$

$C(1,B) = \{X\}$
 $C(2,B) = \text{the empty set}$
 $C(3,B) = \text{the empty set}$
 $C(4,B) = \{Y, Y \rightarrow -X\}$

Here speaker's commitments are not identical because the act of doubting an assertion removes the doubter's commitment to what is doubted. This, then, is one method whereby a speaker can himself withdraw his commitment from a sentence which he has not actually asserted, should he want to. Thus the rule (2) of assertion which commits everyone to an assertion can be seen to be only temporary and by no means permanently binding in its effects. As stated, it does however force the speaker to act in a way relevant to that commitment and thus to engage in argument. In the case of the four step formal argument ' $Y \rightarrow -X$ ' is an implied commitment.

Withdrawal. "W".

- 1 If $W(S)$ then $(\exists t)(S = A(t))$ or $(S = D(t))$.
- 2 If $(\forall n)(\exists i)((n), i A(t))$ or $(\forall n)(\exists i)((n), i D(t))$ and If $(n+1)$
 $W(A(t))$ or $W(D(t))$ then $(\forall i) t$ is not a member of $C(n+1, i)$.

Withdrawal applies only to assertions and doubts. S may be withdrawn only if S occurs earlier in the dialogue and has not already been withdrawn. Intuitively, it makes no sense to withdraw commitment to something which you were not actually committed to in the first place. Further, challenges can only be met, they cannot be withdrawn. In addition, it ought, I think, to be possible to argue someone out of a position of doubt, something along these lines has characteristically been claimed on behalf of criteria in various versions of varying strengths and it therefore seems legitimate to include a mechanism which

at least allows for that possibility. The act of withdrawal removes only the act withdrawn from the commitment-store of the speaker of the act, ie. it does not automatically remove implied commitments featuring the subject of the withdrawn act as antecedent. This prevents the automatic withdrawal of commitments which might have got in as implied commitments but which might also be commitments for some other reason. I leave implied commitments featuring withdrawals as antecedents for possible later challenges. The point of withdrawal is reflexive, ie. it allows a speaker to remove his commitment to one of his own acts.

Doubt. "D".

1 $(\forall n)(\forall i)$ If $((n), i D(S)$ then $(\exists t)(t \in L \text{ and } S = A(t))$.

This states the condition of well-formedness for expressions involving the doubt operator. The commitment effect is basically one of deletion from the doubter's own store and is formally stated as follows:

2 $(\forall n)(\forall i)$ If $((n), i D(A(t))$ then t is not a member of $C(n+1, i)$.

Doubt applies only to sentences in assertions and implied commitments. The only way a doubt can be opened is after an assertion. The primary role of doubt is in the defence of claims and in strategies to that end. A speaker can doubt t only if some speaker is committed to t . The act of doubting an assertion removes the doubter's commitment only to that which is doubted. I do not allow doubt to also remove commitment to conditionals featuring t as antecedent. It is surely reasonable to accept a conditional while denying that its antecedent is actually the case. Classically, of course, $\neg P \vdash P \rightarrow Q$.

Challenge. "?".

1 $(\forall n)(\forall i)$ If $(n) i : ?(S)$ then $(\exists t)(S = (A(t))$ or $S = (D(t))$.

2 $(\forall n)(\forall i)$ If $(n) i : ?(S)$ then $C(n,i) = C(n-1,i)$.

Without Parameters:

1 Assertions: $(\forall n)(\forall i)$ If $(n) i : ?(S)$ Then $(n+1) i$ either: $W(S)$ or: $A(S')$.

2 Doubts: $(\forall n)(\forall i)$ If $(n) i : ?(D(S))$ Then $(n+1) i$ either: $W(D(S))$ or: $A(S')$.

With Parameters:

1 $(\forall n)(\forall i)$ If $(n) i : ?(A(S))/ (Sk...Sn)$ then $(\forall i) C(n,i) = C(n-1,i)$.

2 $(\forall n)(\forall i)$ If $(n) i : ?(D(S))/ (Sk...Sn)$ then $(\forall i) C(n,i) = C(n-1,i)$.

In general, challenge applies to assertion, doubt and implied commitment and has no effect upon either speaker's commitment-store. Intuitively, a challenge does not actually commit one to asserting or denying anything, especially given its use as a basic "why (?)"-type question. Something may be challenged only if some speaker is committed to what is being challenged. Again, it seems counterintuitive to allow speakers to ask for a reason for asserting something which no one has actually asserted.

Challenges are of two types: with parameters and without parameters. Without parameters a challenged doubt must be followed either by a ground for that doubt, in the form of an assertion, or by withdrawal of that doubt if it cannot be grounded. This is a vital feature of the dialogue in so far as it contributes to reflecting in the dialogue my account of the logic of criteria as involving a distinction between the imaginability of the negation of P and the existence of a ground to doubt that P. Clearly, if in a given discourse someone will not assert that P

just because of the imaginability of the negation of P he will fail to assert under just those circumstances which provided the basis of his education in the use of the word. In other words, to doubt where one cannot ground one's doubt in anything in the context is not itself a new strategy in the game, rather it is simply a refusal to play the game at all. The point of a criterion as a basis for warranted assertion is that if the criterion is satisfied then there exists a (defeasible) ground to make that criterial assertion. If, in these circumstances, one still does not want to so assert then unless one is simply rejecting one's training in the use of the assertion one must have some account of how this situation differs from those "paradigmatically suitable" cases which formed the basis of one's original training. The logical advantage of a criterion against the sceptic consists in this mechanism for transferring the burden of proof back to the sceptic. Therefore this feature, the dialectical transfer of the burden of proof, must have a dialogical counterpart. Again, challenge without parameters to an assertion must be followed either by an assertion as ground or by withdrawal of the challenged assertion.

I have given above the two general forms of challenge with parameters. In each case the challenge alleges a conditional ie.1 "I challenge your commitment to S given your commitments to Sk...Sn". Here the conditional $Sk...Sn \rightarrow -S$ is implied by the act. 2 "I challenge your doubt that S given your commitments to Sk...Sn". Here $Sk...Sn \rightarrow S$ is implied by the act. These conditionals need not be featured in the challenger's commitment-store for, as far as possible, merely challenging leaves commitments unaffected. In general, a speaker may challenge with parameters only if the assertion or doubt is an act of the hearer AND the parameters and implied conditional are all commitments of the hearer.

Therefore, challenge with parameters is an appeal to inconsistent commitments.

I hope to offer some characterisation of the differences between criterial and non-criterial conditionals by imposing restrictions on the permissible responses to a challenge made with reference to parameters. That is to say, I will contrast conditionals which reflect criterial relations with what I shall call 'inductively-based' conditionals. These non-criterial conditionals are not, strictly speaking, inductive generalisations or law-like statements traditionally associated, for better or worse, with scientific reasoning. Rather, I have in mind that class of evidence-based assertions which are not properly criterial assertions. For instance, 'the cockerel crowed therefore it must be dawn'. The conclusion is plainly based on evidence and yet that evidence is equally plainly not criterial evidence. The distinction I make is, intuitively, that while in certain cases we are bound to accept certain criterial assertions and cannot reasonably doubt those assertions, we are not so compelled to accept the conclusions of evidence based arguments of the type I have outlined. The distinction must however be manifest in the use of each type of assertion and must therefore flow from the rules of the game. In the first place then, as regards my 'inductively-based' assertions responses to a challenge with parameters must be either, 1 Withdraw the act challenged, 2 Doubt that $Sk...Sn$ obtain or 3 Doubt the implied conditional ie. argue that $Sk...Sn$ may hold but S does not. If a doubt of the form 2 or 3 cannot be grounded then 1 Withdrawal of the act must follow. The criterial version of the rule differs only in the third case. Where we are not compelled by sheer logic, or anything else, to accept the inductively-based conclusion that "it must be dawn" we are

entitled to doubt the conditional. But this cannot be the case as regards criteria. It is not the case that one doubts that the criterial relation holds. For to refuse to accept those relations is just to reject the language. Even in cases of pretence the pretending does not undermine the criterial relation, rather it exploits it. Therefore, I do not allow speakers to doubt conditionals reflecting criterial conditionals. The proper response will be to exhibit a defeater for the assertion made in the present context. Thus the first two options are identical but in the third the inductive argument may be doubted such that one cannot be compelled to accept it, while in the criterial case, that X and Y are criterial relative to Z, cannot be doubted, even though it might be defeated in any context.

For challenge with parameters to assertions I give the following formal rule:

1 $(\forall n)(\forall i)$ If $(n) i : ?(A(S))/ (Sk...Sn)$ then $(n+1) i$ either : $D(Sk)$ or : $D(Sk...Sn) \rightarrow -S$ or : $W(A(S))$.

Because, on my account, there are no negative criteria this use of challenge with parameters will be common to both cases. Further, I note that substituting $-S$ for S and S for $Sk...Sn$ the implied conditional would be $S \rightarrow --S$. There is nothing to prevent logical rules occurring as implied conditionals in this way and even entailments might be so included. Here it should be quite possible for a challenge with parameters to rest upon appeal to a logical proof of the conclusion from the parameters as premises. If there were no proof or the proof could be shown to be defective then it might be stipulated that the challenge be ineffective and inconsequential. I discuss the issue of the introduction of logic to the dialogue system in more detail below but it can already be seen that there is nothing to prevent logical rules occurring as

implied conditionals in this way. If the implied conditional was taken to be an entailment its status could be reflected in the fact that as entailment is not a defeasible relation there will never be a defeater available for it. Hence, doubt of an entailment could never be grounded just because nothing would be allowed to count as a defeater of it.

For challenge with parameters to doubts I give the following rule;
$$2 \ (\forall n)(\forall i) \text{ If } (n) \ i : ?(D(S)) / (Sk \dots Sn) \text{ Then } (n+1) \ i \text{ either } : D(Sk) \text{ or } : D((Sk \dots Sn) \rightarrow S) \text{ or } : W(D(S)).$$

Obviously, it is this version of the rule which governs the inductively-based arguments I referred to earlier. To get the criterial version I substitute an exhibition option for the second doubt option. The rest remains unchanged. Again, logical rules may feature as implied conditionals. While there could not be a defeasible instance of an entailment every instance of a criterial logical rule will be defeasible. That is just to say that while the former is not a defeasible relation the latter is. However, it may be the case in a given context that a criterial inference is actually undefeated in the sense that there is no available ground among the set of evidences with which to make good a doubt. But the conclusion of that inference is still not entailed for that is to say that nothing could defeat it while here the strongest claim made is only that in this case it is actually undefeated. Certainly it is defeasible and defeaters of it exist, but if none are present in a specified context then it is undefeated in that context. In saying that a relation is defeasible I am not contending that it is therefore invariably defeated. Even if critics might hope that were the case.

The basic function of E is as a mechanism for deleting commitments contradictory to any exhibited T, ie. if T can be exhibited then -T is automatically deleted from commitment-stores. Only members of the set of criteria and defeaters may be exhibited.

1 $(\forall n)(\forall i)$ If $(n) i : E(T)$ then $T \in \psi$.

2 $(\forall n)(\forall i)$ If $(n) i : E(T)$ then $(\forall n)(\forall i)$ -T is not a member of $C(n+1, i)$.

Exhibition is always of members of the set of descriptions of criteria and defeaters to which speakers are taken to be mutually agreed. As a precondition of the dialogue speakers agree in the use of this language, although it is by no means any sort of private language. Basically, exhibition highlights a contradiction among the commitments of any speaker committed to -T and automatically removes -T, replacing it with T. Any speaker in the position of possessing contradictory commitments, in effect, loses the dialogue. The act of exhibiting is tantamount to a reminder of the rules of the game. In this it is not dissimilar to that subset of the set of procedural rules which Hamblin calls: 'points of order'. As such "E" cannot be doubted or challenged. That is, I take it that both speakers are speaking the same language and that they will at least agree to descriptions of obvious facts. It is hoped that rule E, which presupposes mutual agreement to the evidence, will overcome the problem of the defeater which is not cited. That is, it may seem that speakers could conclude in favour of a criterial assertion through simply failing to cite a relevant defeater present in that context. However, I take all the evidence to be available to speakers and take speakers to agree on what that evidence is together with the rules for its exploitation. Part of the function of rule E is to ensure that speakers can enforce that implicit agreement by making available to

speakers the machinery to turn presupposed mutual agreements into explicit commitments.

Two final distinctions must be made as regards criteria within the dialogue system. Firstly, criteria will be distinguished from entailments in that while no entailed consequent will be withdrawable at any point in any dialogue it will always be possible to defeat a criterially grounded assertion thus forcing withdrawal of that assertion. Secondly, while both proper inductive generalisations or law-like assertions and criterially grounded assertions can be withdrawn the former will be falsified on production of a counter-instance while a defeater leaves the status of the criterial rule untouched. Thus in the former case we deny the truth of the assertion while in the latter the claim is only withdrawn. Negations of criterial assertions will follow not merely from the obtaining of a defeater but from the criterial assertion of a contrary, where the contrary is established via conceptual interconnections. That is, read purely criterially, $A \rightarrow \neg B$ would not be a wff in the object-language; though it might be treated as such in the meta-language where criterial contraries might be defined and expressed in such a manner (see above, Ch. 6, some remarks on negation following Gabbay's formal representation of non-monotonicity).

These remarks can do no more than sketch a fully-fledged dialogic and the production of any such comprehensive, interacting network of rule-governed systems encapsulating criteria, entailment and induction remains outwith the bounds of this thesis, though there seems no reason why such a project should not be realisable. However, in any dialogical system the admissible moves in the dialogic must define the relations involved. Therefore, if the consequent is not withdrawable at all the relation expressed is that of entailment. If the consequent can only be withdrawn

then the relation is criterial. Finally, where the consequent can be denied and its negation asserted the relation is inductive. For a given conditional then, say, $A \rightarrow B$, if a further piece of evidence is cited such that both A can be denied and $A \rightarrow B$ can be denied then either the conditional is a material implication or a generalized material implication. I propose representing properly inductive law-like conditionals as generalized material implications ie. $(\forall x)(Ax \rightarrow Bx)$. The distinctive feature of the criterial case would be that the conditional, if criterial, could not be denied.

Of course, this is not the only way in which formal logic and inductive inference can be represented in the dialogue. As opposed to this rather integrated conception just outlined it might be possible to demarcate distinct "language-games", as it were, of criterial investigation, of Formal Logic etc. That is, it might be rewarding to see the criterial dialogue as a game which is separate from but interacts with the game of formal logic. To this end it seems natural to introduce the turnstile into the formal language as an elementary predicate sign in the spirit of Curry. To argue that logically, A is entailed by B , that A is a theorem or a truth of logic would all become possible options for speakers. Debate of the validity of suggested truths of logic could be carried out according to separate rules defining the distinct activity of Formal Logic. There might be a number of response options to the assertion of a logical truth. For example, the hearer might justifiably demand a proof of the theorem ie. he might challenge the status of the truth within the system in which it is alleged to be a theorem. Alternately, the hearer might challenge the idea that the statement in the formal dialogue is actually of the form of that

theorem. Thirdly, the hearer might challenge from the viewpoint that being a theorem in a formal system has anything to do with what is the case as regards the formal dialogue. Simple logical theorems might be made available to speakers in the following way, as is illustrated for $P \rightarrow P$:

(1) A: (A(P))	Rule of Assertion
(2) B: ?(A(P))	? -parameters
(3) A: (A(P))	Rule of Assertion
(4) B: ? (P \rightarrow P)	? -parameters
(5) A: $\vdash P \rightarrow P$	\vdash introduction

Commitment-effects would be as follows:

C(1,A) = {P},	C(1,B) = {P}
C(2,A) = {P},	C(2,B) = {P}
C(3,A) = {P, P \rightarrow P},	C(3,B) = {P, P \rightarrow P}
C(4,A) = {P, P \rightarrow P},	C(4,B) = {P, P \rightarrow P}
C(5,A) = {P, P \rightarrow P},	C(5,B) = {P, P \rightarrow P}

The debate might be concluded according to the rules perhaps by invoking a schemata of which A has cited an instance.

It could be argued however that this particular theorem, $\vdash P \rightarrow P$, far from illustrating the possibility of harmonious co-existence and interaction between dialogic and classical formal logic in fact illustrates a fundamental conflict. Jim Mackenzie, for example, in his "No Logic Before Friday" [1985] maintains that, in the context of dialogue, P, so P, is fallacious:

"All instances of P, so P are valid in any standard system of logical inference; but anybody who used the premiss of an argument of this form as a ground for its conclusion would obviously fail to argue successfully. A valid argument form every instance of which is intuitively fallacious is an anomaly for logic. It is only when logic is restored to its dialogical context, and seen to be wider than merely the theory of inference, that the fallacy can be understood as a .. part of a theory of the nature of argument."

From a dialogical viewpoint the prohibition against 'P, so P' is an obvious condition of non-circularity and un informativeness preventing

vicious regresses. In general, Mackenzie's [1985] argument is that:

"For the logician..consideration of dialogue cannot be merely a specialised topic within the discipline like the study of phobias in psychology. Rather the study of dialogue should be the context within which we consider any logical question."

But that is not the only alternative. If formal logic is made available to speakers within a dialogue as a separate though interacting activity and the challenge options are widened sufficiently to include challenges about lack of relevance or inappropriateness, as I have been suggesting, then we might move from the strong, revisionary stance of Mackenzie [1985] to the idea of dialogic as a forum for logic as 'normative science'; in F.P. Ramsey's famous phrase [P.I., 81]. That is, by allowing interaction between speaker's ordinary practices of, say, criterial reasoning and their intuitions which provide relevant data (the stuff of science) and the normative rules of formal logic prescribing the ways in which speakers ought to reason.

III DOUBT AND GROUNDS: TWO CASES

The rules governing speech acts outlined above are intended to exemplify the key feature of criterial theory that ungrounded doubt is always inconsequential. The upshot of these constraints is that the onus is always with the sceptic to ground his doubt. This ready transfer of the burden of proof is itself a vital feature of criterial theory (the vital feature as regards scepticism, above Chapter 6) and is further facilitated by the availability of challenge without parameters which also demands grounding or retraction. The rules of the dialogue naturally force agreement only after every opportunity to ground the doubt has been made available. If the doubt cannot be grounded the inference must go

through on pain of irrationality ie. on pain of consciously possessing incoherent commitments. Failure to accept the inference then amounts to rejecting the rules of the game. Again this is reflective of the type of criterial theory for which I have been arguing. There are therefore at least two cases here which merit closer attention: that in which doubt cannot be grounded and that in which it can, and is, successfully grounded. Below I present an example of a dialogue in which doubt is not grounded. The section following that examines a case in which doubt has been grounded.

In the case below it is assumed that in fact there is no piece of evidence which could ground a doubt. This is for purposes of rule-illustration. The sceptic has his chance in the case examined in the section following that.

CASE ONE.

(1) A: A(P)	Rule Of Assertion
(2) B: ?(P)	Rule Of ?-parameters
(3) A: A(Q)	Rule Of Assertion
(4) B: D(P)	Rule Of Doubt

The pattern of commitment effects generated so far will be as follows:

C(1,A) = {P}	C(1,B) = {P}
C(2,A) = {P}	C(2,B) = {P}
C(3,A) = {P, Q, Q→P}	C(3,B) = {P, Q, Q→P}
C(4,A) = {P, Q, Q→P}	C(4,B) = {Q, Q→P}

At step 4 B doubts P, which he has already challenged at 2, and the Rule of Doubt removes B's commitment to P. The reflexive role of doubt as regards commitment effects is therefore clearly illustrated. A's response is now to challenge B's doubt, thus transferring the burden of proof to B, the sceptic.

(5) A : ?(D(P))/Q, Q → P	Rule Of ?+parameters
--------------------------	----------------------

$C(5,A) = \{P, Q, Q \rightarrow P\}$

$C(5,B) = \{Q, Q \rightarrow P\}$

Although both sets of commitments are unaffected by A's challenge B now has only three options, by ?+parameters B must either:

(6) B : $W(D(P))$

Rule Of Withdrawal

That is, B withdraws his doubt that P and the dialogue terminates in agreement to the conclusion of the criterial inference, P, with the following coherent (rational) commitments,

$C(6,A) = \{P, Q, Q \rightarrow P\}$

$C(6,B) = \{P, Q, Q \rightarrow P\}$

or,

(6') B : $D(Q)$

Rule of Doubt

(7') A : $?(D(Q))$

Rule of ?-parameters

(8') B : $W(D(Q))$

Rule of Withdrawal

(9') A : $?(D(P))/Q, Q \rightarrow P$

?+parameters

(10') B : $W(D(P))$

Rule of Withdrawal

The alternative strategy above generates the following commitments,

$C(6',A) = \{P, Q, Q \rightarrow P\}$

$C(6',B) = \{Q \rightarrow P\}$

$C(7',A) = \{P, Q, Q \rightarrow P\}$

$C(7',B) = \{Q \rightarrow P\}$

$C(8',A) = \{P, Q, Q \rightarrow P\}$

$C(8',B) = \{Q, Q \rightarrow P\}$

$C(9',A) = \{P, Q, Q \rightarrow P\}$

$C(9',B) = \{Q, Q \rightarrow P\}$

$C(10',A) = \{P, Q, Q \rightarrow P\}$

$C(10',B) = \{P, Q, Q \rightarrow P\}$

In other words, although B doubts Q, A again challenges, B can offer no ground for the doubt and is forced to retract it. Consequently, A reiterates his original challenge and the dialogue again terminates in agreement to the conclusion of the criterial inference, P, with coherent (rational) commitments. Given that the context is a purely criterial one the last option open to B is to exhibit a defeater. First, B must exhibit a piece of evidence to which A will reiterate his challenge by parameters to establish the relevance of that evidence. If it is a piece of defeating evidence it will feature in an override rule. If B cannot produce such a rule he must withdraw his original doubt. We have assumed that there is no such defeater and so B's strategy is a bluff which runs as follows,

(6") B : E(T)	Rule of Exhibition
(7") A : ?(D(P))/Q, Q \rightarrow P	?+par's reapplied
(8") B : W(D(P))	Rule of Withdrawal

This strategy generates the commitment pattern below,

C(6",A) = {P, Q, Q \rightarrow P, T}	C(6",B) = {Q, Q \rightarrow P, T}
C(7",A) = {P, Q, Q \rightarrow P, T}	C(7",B) = {Q, Q \rightarrow P, T}
C(8",A) = {P, Q, Q \rightarrow P, T}	C(8",B) = {P, Q, Q \rightarrow P, T}

Again, B cannot ground his doubt in the face of A's challenge, again A restates his earlier challenge forcing B's agreement until the dialogue terminates in agreement to the conclusion of the criterial inference, P, with coherent (rational) commitments.

As the case above makes clear the dialogue is structurally a tree with three branches from node (5). In the case above each branch of the tree terminates in agreement and coherent commitments because B has failed in every attempt to prevent the inference going through by grounding a doubt of the inference, its antecedent or its consequent which, finally, he is forced to accept. It is worth emphasising here that once the sceptic has doubted an assertion the onus becomes his to demonstrate the ground for that doubt. This feature of the dialogue is intended to represent the transfer of the onus of proof in the case of criterially-based assertions.

IV GROUNDING DOUBT

The dialogue below illustrates a standard method for successfully grounding doubt:

Suppose $C(E \rightarrow P)$ and $T // E \rightarrow P$. That is, suppose that E provides a criterial basis for the assertion that P and that T is a named defeater of the criterially-based assertion that P.

(1) A : A(P)	Rule Of Assertion
(2) B : ?(P)	?-parameters
(3) A : E(E)	Rule of Exhibition
(4) B : D(P)	Rule Of Doubt
(5) A : ?(D(P))/E, E \rightarrow P	?+parameters

The five step argument produces the following commitments,

C(1,A) = {P},	C(1,B) = {P}
C(2,A) = {P},	C(2,B) = {P}
C(3,A) = {P, E, E \rightarrow P},	C(3,B) = {P, E, E \rightarrow P}
C(4,A) = {P, E, E \rightarrow P},	C(4,B) = {E, E \rightarrow P}
C(5,A) = {P, E, E \rightarrow P},	C(5,B) = {E, E \rightarrow P}

In essence, A has backed up his claim that P by pointing to some piece of evidence. B cannot doubt or challenge the evidence just because they are taken to be agreed, at least, to what the evidence is. Therefore, B doubts that P and is subsequently challenged by parameters. B's options are now limited by the responses available from a challenge by parameters. That is, ('6') B:D(E1) by Rule Of Doubt is not a permissible move. B cannot doubt or challenge exhibition of evidence to which he agrees. Further, we have presupposed that the present context is purely criterial, or that only the criterial aspect is of relevance. Therefore, the criterial version of challenge by parameters is in force here. Hence, B cannot doubt A's conditional. That is, B cannot doubt a criterial relation; he can only exhibit a defeater in a context. In other words, he cannot deny a criterial relation he can only say 'look, it is defeated in this context'. It looks as if B must now agree with A, or try another option by ?+par's. So far B has failed to ground his doubt. It is quite possible therefore that this dialogue terminates in agreement. That is, at node (6),

(6) B : W(D(P))	Rule of Withdrawal
-----------------	--------------------

The act of withdrawal would make B's commitments coherent and secure his agreement with A. Doubt would fail to be grounded and the dialogue

would terminate in agreement at node (7). However, B does have a decisive alternative strategy,

(6') B : E(T)	Rule of Exhibition
(7') A : ?(D)(P))/E, E → P	?+parameters
(8') B : E(T//E → P)	Rule of Exhibition

In response to A's challenge by parameters B exhibits another piece of evidence which is intended to counter one of A's parameters. A repeats his challenge to determine which parameter B has countered. Exhibition of a defeater of the criterially-based assertion that P, made by A, grounds the doubt alleged by B. The rule of exhibition will replace anything contrary to it in each commitment store, and will therefore override the criterial rule cited by A, showing that the criterial assertion is in fact defeated in this context. The commitment effects are as follows,

C(6',A) = {P, E, E → P},	C(6',B) = {E, E → P}
C(7',A) = {P, E, E → P},	C(7',B) = {E, E → P}
C(8',A) = {E, T, T//E1 → P},	C(8',B) = {E, T, T//E → P}

In the above example doubt is readily and quickly grounded. But this is only for purposes of illustration. In fact dialogues terminate when no legitimate move is left open to speakers and chains of more complex reasonings are perfectly permissible, though these are, in the case of criteria, ultimately terminable in unchallengeable evidence-reports or exhibitions. Below I give an example of a longer terminable chain in a context which is assumed to be purely criterial ie. where each constructed conditional is supposedly criterial

(1)A : A(P)	Rule of Assertion
(2)B : ?(P)	?-parameters
(3)A : A(Q)	Rule of Assertion
(4)B : ?(Q)	?-parameters
(5)A : A(R)	Rule of Assertion
(6)B : ?(R)	?-parameters
(7)A : A(S)	Rule of Assertion
(8)B : D(P)	Rule of Doubt

(9)A : ?(D(P))/Q,Q → P	?+parameters
(10)B : D(Q)	Rule of Doubt
(11)A : ?(D(Q))/R,R → Q	?+parameters
(12)B : D(R)	Rule of Doubt
(13)A : ?(D(R))/S,S → R	?+parameters

The pattern of corresponding commitment effects runs as follows;

C(1,A) = {P},	C(1,B) = {P}
C(2,A) = {P},	C(2,B) = {P}
C(3,A) = {P,Q,Q→P},	C(3,B) = {P,Q,Q→P}
C(4,A) = {P,Q,Q→P},	C(4,B) = {P,Q,Q→P}
C(5,A) = {P,Q,R,R→Q,Q→P},	C(5,B) = {P,Q,R,R→Q,Q→P}
C(6,A) = {P,Q,R,R→Q,Q→P},	C(6,B) = {P,Q,R,R→Q,Q→P}
C(7,A) = {P,Q,R,S,S→R,R→Q,Q→P},	C(7,B) = {P,Q,R,S,S→R,R→Q,Q→P}
C(8,A) = {P,Q,R,S,S→R,R→Q,Q→P},	C(8,B) = {Q,R,S,S→R,R→Q,Q→P}
C(9,A) = {P,Q,R,S,S→R,R→Q,Q→P},	C(9,B) = {Q,R,S,S→R,R→Q,Q→P}
C(10,A) = {P,Q,R,S,S→R,R→Q,Q→P},	C(10,B) = {R,S,S→R,R→Q,Q→P}
C(11,A) = {P,Q,R,S,S→R,R→Q,Q→P},	C(11,B) = {R,S,S→R,R→Q,Q→P}
C(12,A) = {P,Q,R,S,S→R,R→Q,Q→P},	C(12,B) = {S,S→R,R→Q,Q→P}
C(13,A) = {P,Q,R,S,S→R,R→Q,Q→P},	C(13,B) = {S,S→R,R→Q,Q→P}

In the section of dialogue so far examined it is clear that as B stops challenging and starts doubting A's assertions the burden of proof moves from A to B: as would be expected in any reasonable model of the theory of criteria. The reflexive role of the use of the doubt operator is clearly visible in the eliminating effects on commitments which each use of the rule has for the user, namely B, at each point of use, namely, nodes 8, 10 and 12. Therefore, while B can continue to doubt in response to A's challenges he cannot pursue such a strategy indefinitely without being forced to actually ground one of his doubts or withdraw that doubt. Let us suppose that B does continue to doubt.

(14)B : D(S)	Rule of Doubt
(15)A : ?(D(S))	?-parameters

Commitment effects are:

C(14,A) = {P,Q,R,S,S→R,R→Q,Q→P},	C(14,B) = {S→R,R→Q,Q→P}
C(15,A) = {P,Q,R,S,S→R,R→Q,Q→P},	C(15,B) = {S→R,R→Q,Q→P}

B must now try to ground his doubt that (S) in accordance with the rule for challenge without parameters. That is, he must ground his doubt

by appeal to some X such that [if X then -S] or simply withdraw that doubt. A ?-pars cannot be immediately followed by a doubt. In addition B retains his commitment to the conditionals whose antecedents he has been doubting. Intuitively, this is because it is quite possible that the conditional be true while the antecedent is false and because we have assumed a purely criterial context in which the arguments constructed reflect criterial relations to which neither speaker can become uncommitted. Suppose that B cannot ground his doubt that S. This need not yet imply that A has won.

(16)B : W(D(S))	Rule of Withdrawal
(17)A : ?(D(R))/S, S \rightarrow R	?+parameters
(18)B : E(E1)	Rule Of Exhibition
(19)A : ?(D(R))/S, S \rightarrow R	?+parameters

Commitment effects are:

C(16,A) = {P,Q,R,S,S \rightarrow R,R \rightarrow Q,Q \rightarrow P},	C(16,B) = {S,S \rightarrow R,R \rightarrow Q,Q \rightarrow P}
C(17,A) = {P,Q,R,S,S \rightarrow R,R \rightarrow Q,Q \rightarrow P},	C(17,B) = {S,S \rightarrow R,R \rightarrow Q,Q \rightarrow P}
C(18,A) = {P,Q,R,S,S \rightarrow R,R \rightarrow Q,Q \rightarrow P,E1},	C(18,B) = {S,S \rightarrow R,R \rightarrow Q,Q \rightarrow P,E1}
C(19,A) = {P,Q,R,S,S \rightarrow R,R \rightarrow Q,Q \rightarrow P,E1},	C(19,B) = {S,S \rightarrow R,R \rightarrow Q,Q \rightarrow P,E1}

In response to B withdrawing his doubt that (S) A repeats his challenge by parameters forcing B to either withdraw his doubt that R or exhibit some piece of defeating evidence which overrides A's conditional (S \rightarrow R), ie. a defeater of R in this context. B chooses to exhibit a piece of evidence. A repeats his challenge forcing B either to withdraw his doubt that R or to exhibit (E1//S \rightarrow R) from the set of criteria and defeaters. If we suppose that B can so exhibit, that is, that such a piece of defeating evidence is a member of the relevant set, then this branch of the tree will die at node (20) and by the Rule Of Exhibition both commitment-sets will contain the cited defeater and B's commitment to (R), the defeated assertion, will be removed. In other words B will, ultimately, have grounded his doubt that R at node 12.

(20)B : E(E1//S→R)

Rule of Exhibition

However, if no such defeater exists B's move at node 18 will amount to no more than a bluff of the type outlined in the first dialogue. In that case, B's move at the, alternative, node 20 would simply be,

(20')B : W(D(R))

Rule of Withdrawal

This move would, in effect, put R back into B's commitment-set. Consequently, it would allow A to challenge again B's doubt that Q, given his commitment to R and $R \rightarrow Q$, as is illustrated below,

(21')A : ?(D(Q)/R, $R \rightarrow Q$)

?+parameters

In a sense this would just be to return to an earlier node, namely, node 11, where A initially challenged B's doubt in this way, except that, of course, B could not proceed by doubting R, as he did at node 12, for that doubt cannot be grounded. B must therefore attempt to exhibit a defeater of Q, on pain of having to withdraw his doubt that Q. If we again assume that there is in fact no defeater of Q, then B's move must be,

(22')B : W(D(Q))

Rule of Withdrawal

At this stage commitments will be as follows,

$C(22', A) = \{P, Q, R, S, S \rightarrow R, R \rightarrow Q, Q \rightarrow P, E1\}$ $C(22', B) = \{Q, R, S, S \rightarrow R, R \rightarrow Q, Q \rightarrow P, E1\}$

In other words, commitments will differ only in so far as B does not accept that P. Again, A can now reiterate his challenge with parameters made originally at node (9),

(23')A : ?(D(P))/Q, $Q \rightarrow P$

?+parameters

If B cannot exhibit a defeater of P from the set of criteria and defeaters then given that he has failed to ground his doubt that Q he must withdraw his doubt that P,

(24')B : W(D(P))

Rule of Withdrawal

The dialogue would thus terminate in agreement with identical

commitments, as below,

$C(24', A \& B) = \{P, Q, R, S, S \rightarrow R, R \rightarrow Q, Q \rightarrow P, E1\}$

I hope that the dialogue can be seen to be very much a matter of exploring lines of reasoning ie. the developing branches from a given node and frequently returning to that node to reason along other lines in the light of subsequent discoveries. In fact the idea here is a very schematic one. The basic dialogic format allows for constructions which halt easily, which never halt and for a multiplicity in between. Thus formal dialogue has an essential flexibility which is vital to its role as an analogue of actual discourse and its use in modelling the vast range of possible situations.

The ideal system of dialogic, as stated, would be one which represented and contrasted inductive, deductive and criterial arguments and displayed how those methods might interact within a dialogical framework. In addition, it should clearly be the case that the logical character of each class of inference-ticket is exemplified in the rules of the game. But this is plainly an enormous task and while it might make an interesting research programme I can do no more than hint at the way ahead in the present context. However, one issue which I will address here is the problem of distinguishing a criterial basis for assertions from other types of evidential bases. I have suggested earlier that arguments such as 'the cockerel is crowing, therefore it is dawn', or 'he wore a black armband, so someone must have died' might be distinguished from criterially-based assertions in that while there are contexts in which we are forced to accept criterially-based assertions, on pain of rejecting the language we use, this is plainly not so as regards those other, symptomatic, arguments. Again, I have suggested that we cannot be

forced to accept the conclusion of any such argument and have suggested allowing speakers, via challenge with parameters, to doubt the inference-ticket involved in such arguments. Of course, this would make any such purely symptomatic dialogue hopelessly inconclusive, as it perhaps ought to be, but my purpose remains the illustration of criterial reasoning in the first instance and the symptomatic arguments are only of any interest by way of contrast with that type of reasoning. I would suggest that dialogues containing both types of argument are of most interest when construed as exemplifying different ways of getting to the same conclusion. For that purpose I include a certain amount of symptomatic evidence along with the criteria and defeaters and mark it so as to distinguish criterial conditionals from symptomatic conditionals.

Suppose then that we have evidence T which is criterial relative to the assertion that P . Further, suppose that the set ψ of evidence is widened to include evidence ' s ' which is symptomatic relative to P and that the sentence S in ϕ reports ' s '. Again, suppose we also have the defeater E , such that $E//R \rightarrow P$. Consider the following dialogue;

(1)A : A(P)	Rule of Assertion
(2)B : ?(P)	?-parameters
(3)A : A(S)	Rule of Assertion

The commitment effects will be,

$C(1,A) = \{P\},$	$C(1,B) = \{P\}$
$C(2,A) = \{P\},$	$C(2,B) = \{P\}$
$C(3,A) = \{P,S,S \rightarrow P\},$	$C(3,B) = \{P,S,S \rightarrow P\}$

Both speakers are now committed to the symptomatic argument presented by A. But suppose,

(4)B : D(P)	Rule of Doubt
(5)A : ?(D(P))/S, S \rightarrow P	?+parameters

B has doubted that P , removing his commitment to it and in so doing has been challenged by parameters. It is quite open to B now to doubt

that S. I suppose he does as below,

(6)B : D(S)

Rule of Doubt

(7)A : E(s)

Rule of Exhibition

A has responded however by exhibiting 's'. B cannot challenge or doubt the exhibition of evidence. Unlike the criterial case however B may doubt the alleged conditional,

(8)B : (S \rightarrow P)

Rule of Doubt

B is now committed only to 's' and A cannot force B to accept his conclusion. Therefore, this branch from node (1) has died in stalemate. It is perfectly possible to include properly inductive generalisations, represented as $(\forall x)(Ax \rightarrow Bx)$, in a similar manner. In these cases however we might choose to include a counter instance among the evidence. Exhibition of the counter instance could then result in a denial, an assertion of the negation, of that conditional. Although the conclusions would always be negative this would at least allow for more productive dialogues featuring inductive arguments. In contexts in which no counter instance is featured however stalemate would again be the result. This fact reflects the point that while, under certain circumstances, we might be compelled to accept a criterially-based assertion, speakers cannot, I argue, be forced to accept the conclusions of symptomatic or inductive arguments.

Having outlined the formal system I will now discuss some conditions which might be imposed upon it to ensure its coherence and consistency. In any given dialogue mechanisms for dealing with any contradictions which might arise or come to be commitments of a speaker do, of course, already exist. Firstly, contradictory commitments can be challenged, with or without parameters. Secondly, rule 'E' automatically deletes

those commitments in either store which contradict that which has been exhibited. Finally, speakers might easily be allowed to appeal to a theorem, perhaps an Intuitionistic one, to demonstrate the falsity of contradictions. I think it is clear then that while a speaker may become committed to a contradiction the means will exist to prevent that speaker maintaining a contradiction. This fits quite well with the intuition that while any speaker may unwittingly hold a contradiction he must be prepared, as a rational agent, to retract and resolve the contradiction once it is highlighted to him. An obvious candidate to provide consistency of the rules, in a very weak sense, would be the condition that it never be the case that one and the same act is both prohibited and permitted. However, the dialogue is intended to be an analogue of proof and, just as it is part of what a proof is, or part of what "proof" means, that the rules do not prohibit and permit the same move, so that feature should flow from what a dialogue is.

It is important here to say something about consistency over a number of dialogues rather than within any given dialogue. I would argue in this case that it should never be possible to get contradictory conclusions as outputs from two dialogues with the same original inputs. That is, given the same input for any two dialogues, it should never be possible at the last line, as it were, to have one and the same speaker committed to A at the end of one and $\neg A$ at the end of the other. Although I do not have a proof that this is so I think that the condition can be seen to have been met once it is spelled out more fully. Where the members of ψ are identical in any two criterial dialogues it ought not to be possible that at the conclusion of the dialogues one and the same speaker could be committed to an atomic member of ψ in one dialogue and the negation of that atomic member of ψ

in the other dialogue. Intuitively, there seem to be three possibilities as regards criterial dialogues. Firstly, where there are no defeaters, the criterially based assertion A will end up in the commitment store of each speaker. Secondly, where there are defeaters present, A will not end up in the commitment store of each speaker. Now, there would appear to be a third possibility, namely, where we just can't say whether or not a defeater obtains. This case appears to be undecidable. However, in practice, if a speaker cannot tell whether or not a defeater obtains he will not be able to locate that defeater (if there is one) and therefore will not be able to exhibit a defeater. If no defeater can be exhibited then no doubt of any criterially-based assertion can be grounded. If no doubt can be grounded then the criterially-based assertion must be endorsed by both. This is not to claim that the assertion is thereby true or is known but simply that the logical possibility of defeat is not itself a ground to doubt an assertion. Therefore the third option is neither undecidable, nor, I argue, is my account of it unreasonable. For any two dialogues then with identical sets of criteria and defeaters ψ it must be the case that at the conclusion of the dialogues speakers are committed either to A in both or are not so committed in both according to whether or not there is a defeater, but not to something different in each case. My claim then is that, for two dialogues with identical sets ψ , that input will ensure that either A is or is not a member of the commitment store of each speaker, consistently in both dialogues. Therefore, if, at the conclusion of one dialogue, A is a member of the commitment store of a speaker then, at the conclusion of the other, A is a member of that commitment store.

It is clear that in so far as the rules might be considered purely

as formal rules for manipulating symbols so far can the suggested condition of consistency be considered a purely syntactical one. However, the entire system is very nearly an interpreted one and in reality it is very difficult not to give the dialogues a reading and to appeal to considerations about this following that due to the truth or the meaning of one or the other. This consideration points therefore to the condition being in fact a semantic one and, I hope, an intuitively plausible one. Further, I hope that this is not the only point at which the system in general coincides with our intuitions. In fact, the dialogues also give rise to a highly plausible account of how conditionals are formed in natural dialogue. That is, within the formal dialogue we can easily see how conditionals come to be. If, for example, an initial assertion by A, say, "X is in pain", formally, "P" is subsequently doubted by B:

(n) A : P

(n+1) B : D(P)

(In fact ?P would do as well at n+1).

It will be obvious that A's next move is to invoke a piece of evidence, let us say "X is screaming", formally, "Q", thus:

(n+2) A : Q

It is very natural to take A's moves as conjointly implying the claim that $Q \rightarrow P$, "if X is screaming then X is in pain" just because the second assertion is clearly being given as a reason for the first. Hence, A becomes committed to $Q \rightarrow P$ in asserting Q in response to ?(P).

It is also important to note that ϕ and ψ could, in principle, be construed as infinite sets which would generate infinite dialogues. It might be argued that any such dialogue must be undecidable in that we just could not say whether or not a given criterially-based assertion

would end up in the commitment store of both speakers. Again, however if no defeater could be located in a very finite number of moves doubt could not be grounded and assertion would be forced. Therefore, on the basis of the rules, a good case can be made for the dialogues being finite. Here it seems that what is really required to ensure the impossibility of infinite dialogical regress is a rigorous and strict set of rules although a prohibitive rule to prevent looping in dialogues, in the sense of repeating patterns of discourse, might be made explicit. The most plausible condition, I suggest, would be that no assertion already withdrawn may be re-asserted. Therefore, I argue, the dialogue offers a basic account of how conditionals are formed, namely, as commitments produced by responses to challenges to provide grounds for assertion, they are consequences arising from reasons and justifiers. Formal dialogue illuminates the pattern of their construction perspicuously and it is hardly one which stretches our intuitions at all uncomfortably.

At this point it is worth reflecting upon how much of the character of the criterial relation has been captured by the model. In fact a great deal of the nature and status of the relation is represented.

Firstly, the relations themselves are defined in the rule-language as rules of grammar, precisely as they ought to function in natural discourse. The defeasibility essential to criterial inference is represented in a concrete way by explicitly featuring instances of defeaters in the left hand side of the relevant criterial conditionals. Thus, unlike approaches of the type hinted at in Baker [1974] speakers do not have to contend with problematic infinite lists generated by 'unless' clauses. The notion of defeasibility involved in actual practice and the characterisation of the role and limits of scepticism as regards

critical assertion is also further underpinned by the scope of the doubt operator. Full critical support forces the assertion and in the absence of countervailing data the assertion will be fully warranted and "certain" in the sense that because no ground to doubt is available any doubt raised must be groundless and therefore inconsequential. This is not to say however that the assertion is guaranteed to be true, only that in the absence of contravention and with full support it is, by the rules of the game, just what we say. These stipulations then must go some way towards reflecting the distinction between the imaginability of the negation of P and the availability of grounds for doubting that P, which I have held to be fundamental to the logic of criteria. Further, as all moves in the game are linguistic acts, and all properties of linguistic acts are dialogical in character, it will be the role and function of the critical relation in the dialogue which gives it its status and character. Summarizing then, it is clear that the triumvirate of epistemic role, semantic role and the crucial defeasibility, as I perceive these notions to be of relevance to criteria, are all reflected, to some extent, in the formal analogue.

Despite the fact that the task is plainly an enormous one and that a full and proper explication of the logic of criteria presupposes some solution, ultimately, to the problem of non-monotonicity, I would argue that Formal dialogic can offer a fundamentally appropriate framework for a formalization and mathematization which will illuminate and do justice to the practice of critical reasoning.

Finally, while it may well be possible to doubt anything it cannot be possible to doubt everything ie. the meaning of every term we use. The game of doubting must be taught and learned too. Psychologically, at least, we cannot continually doubt whether we are right in applying words

as we do, if doubt is to be a linguistic phenomenon at all, indeed, if language is to be a phenomenon at all. It seems undeniable to me that there exists a practice of basing assertions on criteria, especially as regards psychological attributions. While that practice may be, philosophically, perplexing and problematic it seems nothing but dogmatic to simply reject that practice as wholly wrong. For better or worse such reasoning cannot be simply dismissed. If the logic of criteria is to remain of interest to philosophy however its proponents would do well, I argue, to choose a dialogical context in which to argue their case.

In the last analysis, if scepticism is applicable it may just be that nothing is in fact sceptic-proof or even that a degree of susceptibility to scepticism is in fact an acid-test of good theorizing. Against this however it is worth pointing out that the dialogic bears to some extent upon the Theory of Rationality. If it is allowed that language precedes any concept of rationality in the sense that learning a language involves learning procedures of argument and criticism, allowed, that is, that rationality is a concept acquired when language is acquired, then in learning a language we must learn what is reasonable in argument and criticism. Now, if in learning a language we come to understand the meanings of certain words and sentences via criteria and understand and use criterial rules of inference then criteria spell out what it is rational to infer.

BIBLIOGRAPHY

Abbreviations used in reference to the works of Wittgenstein are as follows:

- P.I. -Philosophical Investigations. Anscombe. Oxford [1963].
- B1.B -The Blue Book. Oxford [1960].
- B.B -The Brown Book. Oxford [1960].
- OC. -On Certainty. Anscombe/Von Wright. Oxford [1975].
- Z. -Zettel. Anscombe/Von Wright. Oxford [1981].
- L & C . -Lectures and Conversations On Aesthetics, Psychology and Religious Belief. Barrett, ed. Oxford [1978].
- R. -Remarks On Colour. Anscombe. Oxford [1977].
- RPP -Remarks on the Philosophy of Psychology Vols 1&2.
- NFL. -Notes For Lectures. Philosophical Review, Vol 77, No. 3, 1968.

Albritton, Rogers [1970]: "On Wittgenstein's Use of The Term Criterion". Pitcher, ed. "Wittgenstein". p.s 231-251. Macmillan.

Alston, W. "Emotion" in "The Encyclopaedia of Philosophy", Vol.2. Macmillan and Free Press (1967).

Austin, J. [1965]: "How to do Things With Words". Oxford.

Baker, G. [1974]: "Criteria: A New Foundation For Semantics" Ratio. 1974.

Baker, G. [1974]: "Defeasibility and Meaning". Article in "Law, Morality and Society", PMS Hacker/J Raz ed. Oxford.

Baker, G. and Hacker P.M.S [1976]: "Critical Notice". Mind. Vol. 85, 1976.

Baker, G. and Hacker, P.M.S [1986]: "Language, Sense & Nonsense". Basil Blackwell Paperback.
First published 1984, in hardback.

Baker, G. and Hacker, P.M.S. [1986A]: "Rules, Scepticism and Language". Paperback.
First published 1984, in hardback.

Bedford, E. [1965-57]: "Emotions". PAS. Vol.57. 6th May 1957.

Bloor, David [1984]: "Wittgenstein: A Social Theory of Knowledge". Macmillan.

Berkeley [1980]: "Berkeley: Philosophical Works". Introduction to "A Treatise Concerning The Principles of Human Knowledge". p.s 61-129. Everyman.

Chomsky, N. [1980]: "Rules and Representations". Basil Blackwell.

- Churchland, P. [1984]: "Matter and Consciousness". The MIT Press.
- Cook [1969]: "Human Beings" in Winch ed. "Studies In The Philosophy of Wittgenstein" p.s 117-151. Routledge/Kegan Paul.
- Cook [1985]: "The Metaphysics of Wittgenstein's On Certainty". Philosophical Investigations 8:2, April 1985.
- Culler, J. [1976]: "Saussure". Fontana Modern Masters.
- Dennett, D. [1981]: "Brainstorms; philosophical essays on mind and psychology. Hassocks.
- Derrida [1973]: "Différance". in Speech and Phenomena, Garver, ed. Northwestern University Press.
- Descartes, R. [1981]: "The Philosophical Works" rendered into English by E.S. Haldane and G.R.T. Ross. Cambridge.
- Dummett, M. [1963/73]: "Realism"/"Philosophical Basis of Intuitionism". Both in "Truth and Other Enigmas". London.
- Dummett, M. [1976]: "What is a Theory of Meaning ? (II)" in Truth and Meaning, Evans/McDowell, eds. Clarendon Press.
- Dummett, M. [1981]: "Frege; Philosophy of Language". London.
- Dummett, M. [1978]: "Truth and Other Enigmas". London.
- Dummett, M. [1987]: "Reply To Prawitz" in "Michael Dummett, Contributions to Philosophy". B. Taylor, ed. Martinus Nijhoff Publishers.
- Evans, G. [1981]: "Semantic Theory and Tacit Knowledge". Article in "Wittgenstein: to Follow a Rule" Holtzman and Leich ed. Routledge/Kegan Paul.
- Flew [1979]: Preface to "A Dictionary of Philosophy". Pan Reference.
- Gabbay [1981]: "Intuitionistic basis for non-monotonic logic" in Proc. Conf on Automated Deduction, Springer Lecture Notes in Computer Science, No. 6, 260-273.
- Garver [1973]: "Speech and Phenomena". Northwestern University Press.
- Geach, P. [1967]: "Mental Acts, Their Content and Their Objects". Studies in Philosophical Psychology. RKP.
- Gert, B. [1986]: "Wittgenstein's Private Language Arguments". Synthese 68 (p.s 409-439).

- Hacker, P.M.S. [1972]: "Insight and Illusion". esp. p.s 283-319. O.U.P.
- Hacker, P.M.S. [1985]: "Insight and Illusion". Revised Edition.
- Hintikka, J. [1962]: "Cogito, Ergo Sum: Inference or Performance ?"
The Philosophical Review, Vol. LXXI, No. 1,
Jan. 1962, pp. 3-32.
- Hume, David [1985]: "A Treatise Of Human Nature". Book One. Penguin
Classics.
- James, W. [1950]: "Principles of Psychology". New York.
Originally [1891] in two volumes. Macmillan.
London.
- Jenkins, J. [1983]: "Understanding Locke". Edinburgh.
- Kenny, Anthony [1984]: "Wittgenstein On The Nature Of Philosophy".
Reprinted in "The Legacy of Wittgenstein",
A. Kenny [1984]. Basil Blackwell.
- Klein, P. [1981]: "Certainty: A Refutation Of Scepticism".
Harvester Press.
- Kripke, Saul [1982]: "Wittgenstein On Rules and Private Language".
Basil Blackwell.
- Kuhn, T. [1964]: "The Structure of Scientific Revolutions".
Chigaco.
- Lakatos, I. [1978]: "Philosophical Papers". Cambridge Paperback Library.
- Lewis, D. [1973]: "Counterfactuals". Oxford.
- Lewis, P.B. [1980]: "On Experiencing The Meaning Of A Word".
- Llewelyn, J. [1987]: "A Point of Almost Absolute Proximity to Hegel".
in Sallis, ed. "Deconstruction and Philosophy".
Chicago.
- Lycan, Gregory W. [1971]: "Non-Inductive Evidence: Recent Work on
Wittgenstein's Criteria". American
Philosophical Quarterly. 1971. p.s 109-125.
- Lyons, J. [1971]: "Chomsky". Fontana Modern Masters.
- Mackenzie, J. [1985]: "No Logic Before Friday". Synthese 63.
- Malcolm, N. [1971]: "Problems of Mind, Descartes to Wittgenstein".
George Allen & Unwin.
- McDermott, D. & Doyle, J. [1980]: "Non-Monotonic Logic I" in
"Artificial Intelligence" 13. 27-39.
- McDermott, D. [1982]: "Non-Monotonic Logic II: non-monotonic modal

theories", in JACM 29 (1) 33-57.

McDowell [1982]: "Criteria, Defeasibility and Knowledge". The Proceedings of the British Academy, London, Vol. LXVIII.

McGinn, Colin [1984]: "Wittgenstein On Meaning". Aristotelian Society Series Vol.1.

Nietzsche [1973]: "Beyond Good And Evil". Hollingdale, tr.

Popper, K [1963]: "Conjectures and Refutations". London.

Quine, W. [1953]: "Two Dogmas of Empiricism" in "From A Logical Point of View". Harper Torchbooks.

Reichling [1961]: "Principles and Methods of Syntax: Cryptanalytical Formalism", *Lingua* 10 pp. 1-7.

Ryle, G. [1949]: "The Concept of Mind". Peregrine.

Searle, ed [1974]: "Philosophy of Language". Oxford Readings.

Staten, H. [1985]: "Wittgenstein and Derrida". Basil Blackwell.

Turner [1984]: "Logics for Artificial Intelligence". Ellis Horwood.

Waismann [1965]: "Principles of Linguistic Philosophy". Macmillan.

Wellman, C. [1981]: "Wittgenstein's Criteria". *Philosophical Review* 1962. Reprinted in "Wittgenstein and the Problem of Other Minds". H. Morick, ed. 1981.

Winch, P. [1987]: "Wittgenstein: Picture and Representation"/
"Eine Einstellung zur Seele". Both in Winch [1987]:
"Trying to Make Sense".

Wright, C. [1984]: "Second Thoughts About Criteria". *Synthese* 58.

Wright, C. [1980]: "Wittgenstein On The Foundations of Mathematics". Duckworth.

Wright, C. [1981]: "Rule-Following, Objectivity and the Theory of Meaning". Article in "Wittgenstein: to Follow a Rule" Holtzman and Leich ed. Routledge/Kegan Paul.

Wright, C. [1982]: "Anti-Realist Semantics: The Role of Criteria". *Philosophy (Supplement)* 1982.

Wright, C. [1986/7]: "Realism, Meaning and Truth". Basil Blackwell.